Dynamical locality and covariance: What makes a physical theory the same in all spacetimes?

Christopher J. Fewster[1*] and Rainer Verch[2†]

[1] Department of Mathematics, University of York, Heslington, York YO10 5DD, U.K.
[2] Institut für Theoretische Physik, Universität Leipzig, 04009 Leipzig, Germany

May 9, 2012

Abstract. The question of what it means for a theory to describe the same physics on all spacetimes (SPASs) is discussed. As there may be many answers to this question, we isolate a necessary condition, the SPASs property, that should be satisfied by any reasonable notion of SPASs. This requires that if two theories conform to a common notion of SPASs, with one a subtheory of the other, and are isomorphic in some particular spacetime, then they should be isomorphic in all globally hyperbolic spacetimes (of given dimension). The SPASs property is formulated in a functorial setting broad enough to describe general physical theories describing processes in spacetime, subject to very minimal assumptions. By explicit constructions, the full class of locally covariant theories is shown not to satisfy the SPASs property, establishing that there is no notion of SPASs encompassing all such theories. It is also shown that all locally covariant theories obeying the time-slice property possess two local substructures, one kinematical (obtained directly from the functorial structure) and the other dynamical (obtained from a natural form of dynamics, termed relative Cauchy evolution). The covariance properties of relative Cauchy evolution and the kinematic and dynamical substructures are analyzed in detail. Calling local covariant theories dynamically local if their kinematical and dynamical local substructures coincide, it is shown that the class of dynamically local theories fulfills the SPASs property. As an application in quantum field theory, we give a model independent proof of the impossibility of making a covariant choice of preferred state in all spacetimes, for theories obeying dynamical locality together with typical assumptions.

1 Introduction

Terrestrial experiments in particle physics are conducted in a weak gravitational field. To interpret their results in terms of QFT models it is therefore necessary that these models can, in principle, be formulated in curved spacetimes without altering their essential physical content, and that one can study and control the limit of weak gravitational fields. This paper is devoted to the first of these issues: specifically, to understanding what

*E-mail: chris.fewster@york.ac.uk
†E-mail: verch@itp.uni-leipzig.de

requirements should be imposed on a theory formulated on a large class of spacetimes to ensure that the physical content is the same in all cases.

Operational concerns dictate a number of restrictions. Experiments are performed in finite regions of spacetime; local causality [31] requires that these experiments should be insensitive to the geometry in the casual complement of the region concerned. Furthermore, the geometrical description of the theory should not be based on preferred systems of reference.

In the context of quantum field theory in curved spacetime, the requirements mentioned so far are implemented within a framework of locally covariant QFT developed by Brunetti, Fredenhagen and Verch (hereafter abbreviated to BFV) in [9] (see also [50]; antecedents of these ideas may be found, e.g., in [22, 35, 36]). There, a quantum field theory defined on all spacetimes is modelled by a functor between a category of globally hyperbolic manifolds and a category of unital $(C)^*$-algebras. Thus to each spacetime $\boldsymbol{M}$ the theory assigns a $(C)^*$-algebra $\mathscr{A}(\boldsymbol{M})$ which might be an algebra of smeared fields, or of local observables; importantly, to each morphism $\psi : \boldsymbol{M} \to \boldsymbol{N}$ between spacetimes[3] there is a corresponding morphism $\mathscr{A}(\psi) : \mathscr{A}(\boldsymbol{M}) \to \mathscr{A}(\boldsymbol{N})$ of $(C)^*$-algebras, so that $\mathscr{A}(\psi \circ \varphi) = \mathscr{A}(\psi) \circ \mathscr{A}(\varphi)$, and with identity morphisms of spacetimes mapped to identity morphisms of $(C)^*$-algebras.

The BFV approach, which we review and develop in Section 3, has significantly advanced the programme of extending results of flat spacetime QFT to curved spacetimes: particular instances include a spin-statistics theorem [50], analogues of the Reeh-Schlieder theorem [47], superselection theory [11, 12], and the perturbative construction of interacting theories in curved spacetime [7, 32, 33]. Applications to *a priori* bounds on Casimir energy densities [27, 25] and new viewpoints in cosmology [18, 20, 51] have also resulted from this circle of ideas.

Somewhat surprisingly, however, it turns out that one may formulate theories in the BFV framework that (at least intuitively) do *not* represent the same physical content in all spacetimes. We will give specific examples in section 4, although these should be regarded as illustrating the range of pathological behaviour, rather than completely describing it. This raises the questions: (a) can one make precise the sense in which such theories fail to have the same content in all spacetimes, and (b) what additional conditions should be imposed to remedy this shortcoming? While we will not completely resolve these issues, we are able to give a framework in which it may be addressed and at least partly resolved.

A fundamental problem is that it is unclear how the 'physical content' of a theory is to be defined in an axiomatic framework. Even a recourse to a Lagrangian setting does not resolve all the issues: see [26] for examples of covariantly defined Lagrangian field theories that do not represent the same physics in all spacetimes. This being so, it is even harder to make precise, by a purely intensional definition, what it means for this content to 'be the same' in different spacetimes.

Given this situation, it seems advisable to allow that there may be many cogent notions of what it means for a theory to represent the same physics in all spacetimes (often

[3] The morphisms are isometric embeddings, preserving orientation and time-orientation, with causally convex image. See Section 2.

abbreviated as SPASs in this paper).[4] Our first aim is to assert principles that should be obeyed by any notion of SPASs and investigate the consequences. In order to do this, we represent any candidate definition of SPASs by the class of theories that conform to it (i.e., an extensional viewpoint); our principles can therefore be expressed as necessary conditions on a class $\mathfrak{T}$ of theories in order that it can serve as a notion of SPASs. Stated as physical principles, they are:

- S1 Every theory in $\mathfrak{T}$ should be locally covariant.
- S2 If $\mathscr{A}$ and $\mathscr{B}$ are (not necessarily distinct) theories in $\mathfrak{T}$, with $\mathscr{A}$ a subtheory of $\mathscr{B}$, and $\mathscr{A}$ and $\mathscr{B}$ coincide in one spacetime, then they should coincide in all spacetimes.

We do not by any means claim that this is an exhaustive prescription and emphasise again that this is not a definition of any particular notion of SPASs but rather a set of principles that should be obeyed by all reasonable notions. Moreover, the term 'coincide' requires precise definition, which will be given below. However, the two conditions together will turn out to be surprisingly strong.

Implementing these principles mathematically, S1 is exactly implemented in the BFV framework and immediately restricts attention to theories that are covariant functors from the category of globally hyperbolic spacetimes to a category Phys of mathematical objects representing 'the physics'. Principle S2 is new, and can be implemented in the BFV framework as follows: if $\mathscr{A}$ and $\mathscr{B}$ are functors representing locally covariant theories, any natural transformation $\zeta : \mathscr{A} \stackrel{.}{\to} \mathscr{B}$ is interpreted as embedding $\mathscr{A}$ as a subtheory of $\mathscr{B}$. The collection of locally covariant theories becomes a category on adopting such subtheory embeddings as morphisms. We will regard the theories as coinciding in some spacetime $\boldsymbol{M}$ if this embedding is an isomorphism in $\boldsymbol{M}$, in which case ζ is called a *partial isomorphism*; the theories coincide in all spacetimes if this condition holds for all $\boldsymbol{M}$, in which case ζ is a natural isomorphism. Principle S2 is then implemented by requiring that all partial isomorphisms between theories in $\mathfrak{T}$ are in fact isomorphisms. In this paper, we will refer to S2, implemented in this way, as *the SPASs property*; however, as indicated above, the axioms above are not expected to be exhaustive. It is conceivable that S2 should be strengthened, to cover situations in which $\mathscr{A}$ and $\mathscr{B}$ may be regarded as coinciding in one spacetime but without the assumption that one is a subtheory of the other. At present it is not known how to implement this mathematically.

Before proceeding, we wish to emphasise the nature of the subtheory embedding with an example. Consider the quantum field theory of the nonminimally coupled scalar field. The field equations $(\Box + \xi R + m^2)\phi = 0$ are evidently independent of the coupling constant in Ricci-flat spacetimes, and this allows the construction of an obvious isomorphism between the algebras of observables for different values of ξ in such spacetimes. However, this does not extend to give a natural transformation between the theories labelled by distinct ξ: the 'obvious isomorphism' does not qualify as a coincidence of the theories, in our sense, even in Ricci-flat spacetimes. A proof of this is sketched at the end of Sect. 3.4.

[4] In principle we even allow that there might even be *no* such notion.

The main result of Sect. 4 is that the SPASs property does not hold in the category of all locally covariant theories unless Phys has rather trivial content; indeed, one can give pairs of theories (which can be otherwise well-behaved) that cannot satisfy the SPASs property; accordingly there is no common notion of SPASs that can accommodate both theories.[5] This is done by an explicit construction that may provide a useful supply of nonstandard locally covariant theories for other purposes. To give a simple outline of one version of our construction, suppose that Phys is the category of $*$-algebras and suppose that $\mathscr{A}$ is a well-behaved theory. We will show that it is possible to construct nonconstant functions on the category of spacetimes, valued in the natural numbers, that are monotonic in the sense that $\chi(\boldsymbol{M}) \leq \chi(\boldsymbol{N})$ for all pairs of spacetimes linked by a morphism $\boldsymbol{M} \to \boldsymbol{N}$. We then define a new theory $\widetilde{\mathscr{A}}$ on objects by $\widetilde{\mathscr{A}}(\boldsymbol{M}) = \mathscr{A}(\boldsymbol{M})^{\otimes\chi(\boldsymbol{M})}$, where the tensor product is the algebraic tensor product on Alg. To any morphism $\psi : \boldsymbol{M} \to \boldsymbol{N}$, we assign a morphism $\widetilde{\mathscr{A}}(\psi) : \widetilde{\mathscr{A}}(\boldsymbol{M}) \to \widetilde{\mathscr{A}}(\boldsymbol{N})$ given by

$$\widetilde{\mathscr{A}}(\psi)(A) = \mathscr{A}(\psi)^{\otimes\chi(\boldsymbol{M})}(A) \otimes \left(\mathbf{1}_{\mathscr{A}(\boldsymbol{N})}\right)^{\otimes(\chi(\boldsymbol{N})-\chi(\boldsymbol{M}))} .$$

A simple computation shows that $\widetilde{\mathscr{A}}$ is a functor from the category of spacetimes to Alg. To check this, note that

$$\widetilde{\mathscr{A}}(\mathrm{id}_{\boldsymbol{M}}) = \mathscr{A}(\mathrm{id}_{\boldsymbol{M}})^{\otimes\chi(\boldsymbol{M})} = \mathrm{id}_{\mathscr{A}(\boldsymbol{M})}^{\otimes\chi(\boldsymbol{M})} = \mathrm{id}_{\widetilde{\mathscr{A}}(\boldsymbol{M})},$$

and that, if $\boldsymbol{M}_1 \xrightarrow{\psi_1} \boldsymbol{M}_2 \xrightarrow{\psi_2} \boldsymbol{M}_3$ then

$$\begin{aligned}
\widetilde{\mathscr{A}}(\psi_2 \circ \psi_1)(A) &= \mathscr{A}(\psi_2 \circ \psi_1)^{\otimes\chi(\boldsymbol{M}_1)}(A) \otimes \left(\mathbf{1}_{\mathscr{A}(\boldsymbol{M}_3)}\right)^{\otimes(\chi(\boldsymbol{M}_3)-\chi(\boldsymbol{M}_1))} \\
&= \left(\mathscr{A}(\psi_2)^{\otimes\chi(\boldsymbol{M}_1)}(\mathscr{A}(\psi_1)^{\otimes\chi(\boldsymbol{M}_1)}(A))\right) \otimes \mathbf{1}_{\mathscr{A}(\boldsymbol{M}_3)}^{\otimes(\chi(\boldsymbol{M}_2)-\chi(\boldsymbol{M}_1))} \otimes \mathbf{1}_{\mathscr{A}(\boldsymbol{M}_3)}^{\otimes(\chi(\boldsymbol{M}_3)-\chi(\boldsymbol{M}_2))} \\
&= \mathscr{A}(\psi_2)^{\otimes\chi(\boldsymbol{M}_2)}\left(\mathscr{A}(\psi_1)^{\otimes\chi(\boldsymbol{M}_1)}(A) \otimes \mathbf{1}_{\mathscr{A}(\boldsymbol{M}_2)}^{\otimes(\chi(\boldsymbol{M}_2)-\chi(\boldsymbol{M}_1))}\right) \otimes \mathbf{1}_{\mathscr{A}(\boldsymbol{M}_3)}^{\otimes(\chi(\boldsymbol{M}_3)-\chi(\boldsymbol{M}_2))} \\
&= \widetilde{\mathscr{A}}(\psi_2)(\widetilde{\mathscr{A}}(\psi_1)(A)) \\
&= \left(\widetilde{\mathscr{A}}(\psi_2) \circ \widetilde{\mathscr{A}}(\psi_1)\right)(A)
\end{aligned}$$

for any $A \in \mathscr{A}(\boldsymbol{M}_1)$, using the unit-preserving property of Alg-morphisms. Thus the functor $\widetilde{\mathscr{A}}$ satisfies the definition of a locally covariant quantum field theory. However one cannot expect both $\mathscr{A}$ and $\widetilde{\mathscr{A}}$ to have the same physical content in all spacetimes as the theory consists of $\chi(\boldsymbol{M})$ copies of the basic theory $\mathscr{A}(\boldsymbol{M})$ in each spacetime $\boldsymbol{M}$. Developing the example further, if χ has 1 and ℓ as its minimum and maximum values, then there are successive subtheory embeddings $\mathscr{A} \stackrel{.}{\to} \widetilde{\mathscr{A}} \stackrel{.}{\to} \mathscr{A}^{\otimes\ell}$, each of which is a partial isomorphism, but whose composite is not an isomorphism [this is a mild assumption on $\mathscr{A}$]; so at least one of the partial isomorphisms cannot be an isomorphism. Thus the three theories $\{\mathscr{A}, \widetilde{\mathscr{A}}, \mathscr{A}^{\otimes\ell}\}$ cannot conform to a single notion of SPASs, and indeed the same is true of at least one of the pairs $\{\mathscr{A}, \widetilde{\mathscr{A}}\}$ or $\{\mathscr{A}^{\otimes\ell}, \widetilde{\mathscr{A}}\}$. Of course, if $\mathscr{A}$ is a familiar theory

[5] At the end of Sect. 4 we even construct *single* theories that cannot satisfy *any* notion of SPASs.

that one would regard intuitively as representing SPASs, then one would regard $\widetilde{\mathscr{A}}$ as 'obviously' not representing SPASs by the same intuitive standard. Our aim in formalising these questions is to provide a framework in which such judgements can be made without relying on intuition.

This result raises the question as to what conditions might produce a class of theories obeying the SPASs property. Our answer to this involves a closer examination of the specification of the physics associated with local regions in globally hyperbolic spacetimes. One of the attractive features of the BFV framework is that it gives a definition for the local physics associated to a region O in spacetime $\boldsymbol{M}$, essentially by considering the region O (with the geometry restricted from $\boldsymbol{M}$) as a spacetime in its own right. We shall regard this as a kinematical description of local physics. In Section 5, we introduce a new description of the local physics in O that is based on dynamics: the local physics in O is that portion of the physics on the whole spacetime that is invariant under modifications of the spacetime metric in the causal complement of O in $\boldsymbol{M}$. The effect of a modification to the metric is captured by the *relative Cauchy evolution* introduced by BFV, which is closely related to the dynamics of the theory.[6] We investigate the basic properties of the resulting 'dynamical net'; while it has a number of features in common with the 'kinematic net' it lacks others, notably the local covariance property of the kinematic net does not hold for the dynamical net in general.

The situation in which the kinematic and dynamical nets coincide is of particular interest, and those theories for which it holds will be said to be *dynamically local*. As we show in Section 6, dynamically local theories have a number of good properties: they are additive, have good covariance properties for the dynamical net, and (under a mild additional assumption) obey *extended locality* in the sense that the local physics for spacelike separated regions intersect only trivially.[7] The scope for constructing pathological theories of the sort discussed in Section 4 is significantly reduced and even eliminated if the theory has no nontrivial automorphisms (as is expected for a theory of local observables). Moreover, as is shown in Theorem 6.10, the class of dynamically local theories has the SPASs property. Accordingly, the concept of dynamical locality provides a first answer to the problem of isolating those theories that can be regarded as representing the same physics in all spacetimes, and appears to be a useful addition to the axiomatic framework in curved spacetimes.

As an application of these results to QFT, we give the first model independent proof of the impossibility of selecting a single 'natural' state in each spacetime (Section 6.3) for any nontrivial dynamically local theory with the extended locality property, on the assumption that the supposed natural state has the Reeh–Schlieder property in some spacetime. (Here,

[6]The functional derivative of the relative Cauchy evolution with respect to the metric perturbation can be interpreted as the stress-energy tensor, from which viewpoint the relative Cauchy evolution is a replacement for a classical action in this framework.

[7]What 'trivial' means here will depend on the category Phys employed to describe the physics. In the categories of $(C)^*$-algebras employed in QFT, this means that the intersection consists of complex multiples of the algebra unit. Extended locality was originally introduced in [49, 38] in the context of algebraic QFT in Minkowski space.

we say that a theory is trivial if it is equivalent to the theory whose algebra of observables consists of complex multiples of the unit in every spacetime.) Neither of these additional assumptions seem unreasonable; in particular, our result applies to any theory that reduces, in Minkowski space, to a Wightman or Haag–Kastler theory obeying standard conditions and with the natural state reducing to the Minkowski vacuum state. It is worth noting that the SPASs property is used as a technical input to the proof: the given theory is shown to coincide with the trivial theory in one spacetime, and must therefore do so in all.

In addition to these results, and as a necessary technical tool in proving them, we make a thorough study of the relative Cauchy evolution, deepening the investigation begun by BFV. We particularly study the covariance properties of the relative Cauchy evolution, and the way in which subtheory embeddings intertwine the relative Cauchy evolutions of different theories. Our methods, wherever possible, are adapted to the widest possible categorical setting, to emphasise the applicability of underlying ideas; all the key concepts are expressed in terms of universal properties, which makes for efficient proofs that are portable between different physical settings. On the geometrical side, we also adapt and extend the spacetime deformation methods introduced in [29]: in spacetime dimension $n \geq 2$, these techniques allow us to partition the category of spacetimes into connected components labelled by equivalence classes of Riemannian manifolds of dimension $n-1$, modulo orientation-preserving diffeomorphisms. Here, connectedness is understood in terms of the existence of chains of 'Cauchy wedges' from one spacetime to another. One might conjecture that a more detailed study of the category of spacetimes from this viewpoint would give a cohomology theory with many ramifications. Indeed, following our suggestion, Sanders has shown that one may regard various freedoms arising in the construction of the Dirac field in curved spacetimes in precisely such a cohomological way [48]. Appendix A provides a body of material on spacetime structure, required in the body of the paper, particularly in relation to different notions of causal complement. We hope that a number of these developments will be useful for other purposes.

A separate paper [28] is devoted to an investigation of the dynamical locality for various linear theories, both as classical and quantum fields. It is shown that dynamical locality is satisfied by the massive minimally coupled free scalar field. At zero mass, dynamical locality fails; however, this can be understood as an expression of the rigid gauge symmetry of the minimally coupled massless field. When the theory is quantised as a (rather simple) gauge theory, dynamical locality is restored in dimension $n > 2$ (and even in dimension $n = 2$ if one restricts to connected spacetimes). What significance can be read into this special case is currently unclear. Dynamical locality is known to hold for the nonminimally coupled scalar field at any value of the mass [23], and work on other models, including the algebra of Wick products is under way.

2 Categories of spacetimes

We begin by defining the categories of spacetimes that will be used as the arena for locally covariant theories. This serves to fix our notation and terminology; while much of this

material is fairly standard, our study of the connectedness properties of the categories with respect to wedges gives a new viewpoint on classical results on deformations of globally hyperbolic spacetimes [29]. Some of the details are deferred to Appendix A, which also contains a number of useful results on causal structure.

2.1 Globally hyperbolic spacetimes

A *globally hyperbolic spacetime* of dimension n is a quadruple $(\mathcal{M}, \boldsymbol{g}, \mathfrak{o}, \mathfrak{t})$ such that

- $\mathcal{M}$ is a smooth paracompact orientable nonempty n-manifold with finitely many connected components
- $\boldsymbol{g}$ is a smooth time-orientable metric of signature $+ - \cdots -$ on $\mathcal{M}$
- $\mathfrak{o}$ is a choice of orientation, i.e., one of the connected components of the set of nowhere-zero smooth n-form fields on $\mathcal{M}$
- $\mathfrak{t}$ is a choice of time-orientation for $\boldsymbol{g}$, i.e., one of the connected components of the set of nowhere-zero smooth $\boldsymbol{g}$-timelike 1-form fields on $\mathcal{M}$

and such that the resulting causal structure is globally hyperbolic, i.e., there are no closed causal curves and the intersection of the causal past and future of any pair of points is compact.[8] For global hyperbolicity, it is sufficient that $\mathcal{M}$ contains a Cauchy surface [44, Cor. 14.39], that is, a subset met exactly once by every inextendible timelike curve in the spacetime.[9] A Cauchy surface is necessarily a closed achronal topological hypersurface met (at least once) by every inextendible causal curve [44, Lem. 14.29]. All Cauchy surfaces of a given globally hyperbolic spacetime $\boldsymbol{M}$ are homeomorphic [44, Cor. 14.27]. Further, $\boldsymbol{M}$ admits smooth spacelike Cauchy surfaces [3, Thm 1.1]; given any such surface Σ, it is possible to construct a diffeomorphism $\rho : \mathbb{R} \times \Sigma \to \mathcal{M}$ with the following properties (see [5, Thm 1.2] and [4, Thm 2.4]):

- $\rho_0(\cdot) = \rho(0, \cdot)$ is the inclusion $\Sigma \hookrightarrow \boldsymbol{M}$
- for each $t \in \mathbb{R}$, $\rho(\{t\} \times \Sigma)$ is a smooth spacelike Cauchy surface
- $\rho_* \partial/\partial t$ is future-directed
- the pulled back metric splits in the form $\rho^* \boldsymbol{g} = \beta dt \otimes dt - \boldsymbol{h}_t$ where $\beta \in C^\infty(\mathbb{R} \times \Sigma)$ is positive, and $t \mapsto \boldsymbol{h}_t$ is a smooth map into the smooth Riemannian metrics on Σ.

The Cauchy surface Σ has a unique orientation $\mathfrak{w}$ such that $\mathfrak{o} = \mathfrak{t} \wedge \mathfrak{w}$ (extending the wedge product to equivalence classes of forms in an obvious way) and we will regard this as the canonical orientation on Σ. Equipping $\mathbb{R} \times \Sigma$ with the orientation corresponding to $dt \wedge \mathfrak{w}$, the diffeomorphism ρ is promoted to an orientation-preserving diffeomorphism

[8] This appears weaker than the definition given, e.g., in [31], but the two are equivalent by Thm. 3.2 in [6].

[9] A slightly stronger definition is employed in [31], where a Cauchy surface is defined to be an edgeless acausal set intersected (exactly once) by every inextendible causal curve. This equates to an acausal Cauchy surface in our terminology.

(abbreviated as oriented-diffeomorphism); this preserves time-orientations on declaring $\partial/\partial t$ to be future-pointing. We refer to the result of the above construction as the *normal form* for globally hyperbolic spacetimes.

There are, of course, many globally hyperbolic spacetimes.

Proposition 2.1 *Every smooth, paracompact oriented $(n-1)$-manifold that is connected (resp., has finitely many connected components) is oriented-diffeomorphic to a smooth spacelike Cauchy surface of a spacetime in Loc_0 (resp., Loc).*

Proof: In the connected case, suppose an $(n-1)$-manifold Σ is given with orientation defined by a nonvanishing $(n-1)$-form ω. Equip Σ with a complete Riemannian metric $\boldsymbol{h}$ [43] and endow $\mathbb{R}\times\Sigma$ with metric $dt\otimes dt-\boldsymbol{h}$, orientation $dt\wedge\omega$ and time-orientation $\partial/\partial t$. Then the resulting structure is globally hyperbolic with each $\{t\}\times\Sigma$ as a Cauchy surface [34, Prop. 5.2] that is oriented-diffeomorphic to Σ with orientation ω. In the disconnected case, we perform this construction on each connected component and form the union. $\square$

2.2 The categories Loc and Loc_0

The globally hyperbolic spacetimes (of dimension n) form the objects of a category Loc. By definition, a morphism ψ in Loc between $\boldsymbol{M}=(\mathcal{M},\boldsymbol{g},\mathfrak{o},\mathfrak{t})$ and $\boldsymbol{M}'=(\mathcal{M}',\boldsymbol{g}',\mathfrak{o}',\mathfrak{t}')$ is a smooth embedding (also denoted ψ) of $\mathcal{M}$ in $\mathcal{M}'$ whose image is causally convex in $\boldsymbol{M}'$ and such that $\psi^*\boldsymbol{g}'=\boldsymbol{g}$, $\psi^*\mathfrak{o}'=\mathfrak{o}$ and $\psi^*\mathfrak{t}'=\mathfrak{t}$. Thus the embedding is isometric and respects orientation and time-orientation. In particular, any diffeomorphism putting a globally hyperbolic spacetime into normal form is itself an isomorphism in Loc.

Causal convexity of the image entails that every smooth causal curve with ends contained in the image is contained entirely in it. In particular, if O_1 and O_2 are any distinct connected components of the image of $\boldsymbol{M}$ there can be no causal curve joining a point of O_1 to a point of O_2: i.e., $O_1\subset O_2^\perp:=\boldsymbol{N}\setminus J_{\boldsymbol{N}}(O_2)$ and likewise $O_2\subset O_1^\perp$. In fact, as the O_i and hence $J_{\boldsymbol{N}}(O_i)$ are necessarily open (see, e.g., Lem. A.8), we have the slightly stronger condition $O_1\subset O_2':=\boldsymbol{N}\setminus \mathrm{cl}\, J_{\boldsymbol{N}}(O_2)$ and $O_2\subset O_1'$. It is possible, nonetheless, that the closures of O_1 and O_2 can intersect nontrivially. Note that we have introduced two distinct notions of causal complement, both of which will be needed in what follows. Some relations between these two definitions and their various properties are discussed in Appendix A, in which standard definitions of causal structure (such as the set $J_{\boldsymbol{N}}(O)$ just used) are also recalled – see Appendix A.2.

We will also study the full subcategory of Loc with connected spacetimes as objects, which will be denoted Loc_0. Each connected component of an Loc object $\boldsymbol{M}$ is an Loc_0 object; we denote the set of components of $\boldsymbol{M}\in\mathsf{Loc}$ by $\mathrm{Cpts}(\boldsymbol{M})$. Each Loc morphism $\psi:\boldsymbol{M}\to\boldsymbol{N}$ comprises one or more Loc_0 morphisms: to each component $\boldsymbol{B}\in\mathrm{Cpts}(\boldsymbol{M})$ there is a unique component $\boldsymbol{C}\in\mathrm{Cpts}(\boldsymbol{N})$ containing the image $\psi(\boldsymbol{B})$ of $\boldsymbol{B}$, and the restriction of ψ to $\boldsymbol{B}$ yields a Loc_0-morphism $\psi_{\boldsymbol{B}}^{\boldsymbol{C}}:\boldsymbol{B}\to\boldsymbol{C}$. Conversely, any collection of Loc_0-morphisms $(\psi_{\boldsymbol{B}}^{f(\boldsymbol{B})})_{\boldsymbol{B}\in\mathrm{Cpts}(\boldsymbol{M})}$ where $f:\mathrm{Cpts}(\boldsymbol{M})\to\mathrm{Cpts}(\boldsymbol{N})$ and $\psi_{\boldsymbol{B}}^{f(\boldsymbol{B})}:\boldsymbol{B}\to f(\boldsymbol{B})$ defines a Loc morphism, provided that their images are all causally disjoint in the above

sense. It is not required that every component of $\boldsymbol{N}$ should contain the image of one or more components of $\boldsymbol{M}$.

Two particular classes of Loc and Loc_0 morphisms will be used extensively in what follows: *canonical inclusions* and *Cauchy morphisms*. Inclusions arise as follows. For any $\boldsymbol{M}$ in Loc (and hence Loc_0) let $\mathscr{O}(\boldsymbol{M})$ be the set of open globally hyperbolic subsets[10] of $\boldsymbol{M}$ with at most finitely many connected components all of which are mutually causally disjoint, and let $\mathscr{O}_0(\boldsymbol{M})$ be the set of connected open globally hyperbolic subsets of $\boldsymbol{M}$. For each $\boldsymbol{M} = (\mathcal{M}, \boldsymbol{g}, \mathfrak{o}, \mathfrak{t}) \in \mathsf{Loc}$, any nonempty $O \in \mathscr{O}(\boldsymbol{M})$ induces an object $\boldsymbol{M}|_O = (O, \boldsymbol{g}|_O, \mathfrak{o}|_O, \mathfrak{t}|_O)$ of Loc, which we call the *restriction* of $\boldsymbol{M}$ to O, and the subset inclusion of O in $\boldsymbol{M}$ induces a Loc-morphism $\iota_{\boldsymbol{M};O} : \boldsymbol{M}|_O \to \boldsymbol{M}$ that we call a canonical inclusion. Any morphism $\boldsymbol{L} \xrightarrow{\psi} \boldsymbol{M}$ induces a canonical isomorphism $\tilde{\psi} : \boldsymbol{L} \xrightarrow{\cong} \boldsymbol{M}|_{\psi(\boldsymbol{L})}$ so that $\psi = \iota_{\boldsymbol{M};\psi(\boldsymbol{L})} \circ \tilde{\psi}$. If $O \in \mathscr{O}_0(\boldsymbol{M})$ for $\boldsymbol{M} \in \mathsf{Loc}_0$ then $\iota_{\boldsymbol{M};O}$ is also a Loc_0-morphism, provided O is nonempty.

A morphism $\psi : \boldsymbol{M} \to \boldsymbol{N}$ will be described as a *Cauchy morphism*, or simply as *Cauchy* if its image contains a Cauchy surface for $\boldsymbol{N}$. All identity morphisms in Loc_0 and Loc are Cauchy and compositions of Cauchy morphisms are Cauchy (Lem. A.3 in Appendix A), so the globally hyperbolic spacetimes with Cauchy morphisms define subcategories of Loc_0 and Loc. As there are slightly different definitions of Cauchy surface in the literature, of which we have adopted the weakest, the following observation is worth recording (see Appendix A for the proof).

Proposition 2.2 *If $\psi : \boldsymbol{M} \to \boldsymbol{N}$ is Cauchy (in Loc_0 or Loc) then $\psi(\boldsymbol{M})$ contains a Cauchy surface of $\boldsymbol{N}$ that is smooth, spacelike and acausal. Moreover, the Cauchy surfaces of $\boldsymbol{M}$ and $\boldsymbol{N}$ are homeomorphic and their smooth spacelike Cauchy surfaces are oriented-diffeomorphic.*

A key fact for our purposes is that morphisms in Loc_0 whose domain has compact Cauchy surfaces are always Cauchy. The following is an immediate consequence of Prop. A.1 in Appendix A together with Prop. 2.2.

Proposition 2.3 *(a) Suppose $\psi : \boldsymbol{M} \to \boldsymbol{N}$ in Loc_0, where $\boldsymbol{M}$ has compact Cauchy surfaces. Then ψ is Cauchy and the smooth spacelike Cauchy surfaces of $\boldsymbol{N}$ are oriented-diffeomorphic to those of $\boldsymbol{M}$. (b) Suppose $\psi : \boldsymbol{M} \to \boldsymbol{N}$ in Loc and suppose $\boldsymbol{B} \in \mathrm{Cpts}(\boldsymbol{M})$ has compact Cauchy surfaces. If $\boldsymbol{C}$ is the component of $\boldsymbol{N}$ containing $\psi(\boldsymbol{B})$, then $\psi_{\boldsymbol{B}}^{\boldsymbol{C}}$ is Cauchy and $\boldsymbol{C}$ has smooth spacelike Cauchy surfaces oriented-diffeomorphic to those of $\boldsymbol{B}$. Moreover, $\boldsymbol{C}$ cannot contain the image of any component of $\boldsymbol{M}$ other than $\boldsymbol{B}$ (because $\psi(\boldsymbol{B})$ has trivial causal complement in $\boldsymbol{C}$).*

2.3 Deformation arguments and "wedge connectedness"

Globally hyperbolic spacetimes with oriented-diffeomorphic Cauchy surfaces can be deformed into one another, a result going back to [29] (although the emphasis on orientation

[10] See [44, Def. 14.20]. Note that the open globally hyperbolic subsets of a globally hyperbolic spacetime are precisely the open causally convex subsets.

here is new). In the present language, this can be stated as follows:

Proposition 2.4 *Two spacetimes* $\boldsymbol{M}$, $\boldsymbol{N}$ *in* Loc_0 *(resp.,* Loc*) have oriented-diffeomorphic Cauchy surfaces if and only if there exists a chain of Cauchy morphisms in* Loc_0 *(resp.,* Loc*) forming a diagram*

$$\boldsymbol{M} \leftarrow \boldsymbol{F} \rightarrow \boldsymbol{I} \leftarrow \boldsymbol{P} \rightarrow \boldsymbol{N}. \tag{2.1}$$

Proof: If such a chain of Cauchy morphisms exists, then $\boldsymbol{M}$ and $\boldsymbol{N}$ have oriented-diffeomorphic smooth spacelike Cauchy surfaces by Prop. 2.2. The converse is an elaboration of [29, Appx C] and is given for completeness in Appendix A. □.

The chain of morphisms here is far from unique. We will find it useful to regard this result in the following manner. In a general category, a *wedge* is any pair of morphisms with common domain, i.e., a diagram of form $B \xleftarrow{f} A \xrightarrow{g} C$. Proposition 2.4 then asserts that spacetimes with oriented-diffeomorphic Cauchy surfaces are connected by two *Cauchy wedges*, i.e., wedges consisting of Cauchy morphisms. This shows that Loc decomposes into "Cauchy-wedge-connected" components labelled by equivalence classes of Cauchy surfaces modulo oriented-diffeomorphisms; the same is true for Loc_0 on restriction to connected $(n-1)$-manifolds. We remark in passing that some oriented $(n-1)$-manifolds belong to the same equivalence class as their orientation reverse (e.g., $\mathbb{R}^{n-1}$, S^{n-1}) while others (e.g., the three-dimensional lens space $L_5(1,1)$) are not, and are sometimes called *chiral* (see, e.g., [42]). Thus any two spacetimes with Cauchy surfaces diffeomorphic to $\mathbb{R}^{n-1}$ (with the standard differential structure) are linked by a chain of Cauchy morphisms, but spacetimes with inequivalently oriented chiral Cauchy surfaces belong to different Cauchy components of Loc.

We also have another connectedness result, this time for the general class of wedges. To this end, we first introduce the particularly useful class of *diamond* subsets of a globally hyperbolic spacetime following Brunetti and Ruzzi [12]. We will also consider *multi-diamonds*, that is, unions of finitely many causally disjoint diamonds.

Definition 2.5 *Let* $\boldsymbol{M}$ *be a spacetime in* Loc. *A* Cauchy ball *in a Cauchy surface* Σ *of* $\boldsymbol{M}$ *is a subset* $B \subset \Sigma$ *for which there is a chart* (U, ϕ) *of* Σ *such that* $\phi(B)$ *a nonempty open ball in* $\mathbb{R}^{n-1}$ *whose closure is contained in* $\phi(U)$. *A* diamond *in* $\boldsymbol{M}$ *is any open relatively compact subset of the form* $D_{\boldsymbol{M}}(B)$ *where* B *is a Cauchy ball in some Cauchy surface* Σ. *We say that the diamond has base* B *and that it is based on the Cauchy surface* Σ.[11] *A* multi-diamond *is a union of finitely many causally disjoint diamonds, and therefore takes the form* $D_{\boldsymbol{M}}(B)$ *where* B *is a* Cauchy multi-ball, *i.e., a union of finitely many causally disjoint Cauchy balls.*

Using Theorem 4.1 and Remark 4.14 in [5], for any Cauchy multi-ball B there is a (nonunique) Cauchy surface in which it is contained. This observation allows us to extend the properties of diamonds established in [12] to show that, in spacetime dimension $n \geq 3$, any (multi)-diamond is (among other properties) open, relatively compact, simply

[11]Neither the base B nor the Cauchy surface Σ are uniquely associated with the diamond.

connected, and has a nonempty causal complement $O' = \boldsymbol{M} \setminus \mathrm{cl}(J_{\boldsymbol{M}}(O))$, whose intersection with any connected component of $\boldsymbol{M}$ is itself connected. Diamonds are connected. A number of further properties of(multi-)diamonds are given in Appendix A. In particular, in Lemma A.9 we demonstrate for completeness that any (multi-)diamond is causally complete in the sense that $O = O''$.

In what follows, we will say that a spacetime $\boldsymbol{D}$ is a (multi-)diamond if it is isomorphic to a restriction $\boldsymbol{M}|_O$, where O is a (multi-)diamond of some $\boldsymbol{M}$ in Loc or Loc_0. A *truncated (multi-)diamond* will refer to any intersection of a (multi-)diamond with an open globally hyperbolic neighbourhood of a Cauchy surface on which it is based.

Proposition 2.6 *If $\boldsymbol{M}$ and $\boldsymbol{N}$ are any globally hyperbolic spacetimes in Loc (resp., Loc_0) then there exists a chain of (not necessarily Cauchy) morphisms in Loc (resp., Loc_0) creating a diagram of the form* (2.1).

Proof: Let O_1 and O_2 be diamond regions in $\boldsymbol{M}$ and $\boldsymbol{N}$ respectively. The Cauchy surfaces of $\boldsymbol{M}|_{O_1}$ and $\boldsymbol{N}|_{O_2}$ are oriented-diffeomorphic (they are homeomorphic to $\mathbb{R}^{n-1}$), so we may apply Prop. 2.4 to obtain a chain of Cauchy morphisms $\boldsymbol{M}|_{O_1} \leftarrow \boldsymbol{F} \rightarrow \boldsymbol{I} \leftarrow \boldsymbol{P} \rightarrow \boldsymbol{N}|_{O_2}$ and we compose at the two ends with $\iota_{\boldsymbol{M};O_1}$ and $\iota_{\boldsymbol{N};O_2}$ to obtain the required result. $\square$

3 Locally covariant theories

3.1 Categories of physical systems

The focus of BFV was on quantum field theories, described in terms of algebras of observables and suitable state spaces. Here, we wish take a more general approach in order to encompass a broader range of physical theories.

Suppose a certain type of physical system is to be formulated in a locally covariant way on globally hyperbolic spacetimes. We suppose that the physical systems concerned can be represented mathematically by objects of a category Phys, whose morphisms correspond to embeddings of one such system in another.

The general conditions imposed on Phys will be that all its morphisms are monic, that it has equalisers, intersections and unions [in the categorical sense, which do not necessarily coincide with the set-theoretic notions; the relevant definitions are given in Appendix B], and that it possesses an initial object, denoted $\mathcal{I}$ and representing the trivial physical system of the given type, which is uniquely embedded in every system $\mathcal{A}$ via a morphism denoted $\mathcal{I}_{\mathcal{A}}$ (we have $\alpha \circ \mathcal{I}_{\mathcal{A}} = \mathcal{I}_{\mathcal{B}}$ for every $\alpha : \mathcal{A} \to \mathcal{B}$). As general references on category theory, see [41, 1]; our discussion of subobjects and their intersections and unions follows [21].

Possible candidates for the category Phys abound. The BFV setting corresponds to categories such as: (a) the category Alg of unital $*$-algebras with unit preserving faithful $*$-homomorphisms; (b) the category $\mathsf{C^*\text{-}Alg}$ of unital C^*-algebras with unit preserving faithful $*$-homomorphisms; (c) the category TAlg of unital topological $*$-algebras with continuous unit preserving faithful $*$-homomorphisms as morphisms; in each case the initial object $\mathcal{I}$

is the complex number field $\mathbb{C}$ with 1 as the unit, complex conjugation as the $*$-operation and additional topological structure as appropriate to the category concerned. Elsewhere, we will discuss a category Sys, whose objects are $*$-algebras or C^*-algebras together with a suitable subset of the states thereon. More widely, our discussion could also be applied to classical mechanical or field systems – the use of a general category Phys emphasises these possibilities. As a classical example, Phys could be the category of presymplectic vector spaces with injective symplectic linear maps as morphisms, and the trivial symplectic space as the initial object.

The categorical notions mentioned above can be illustrated easily in $\mathsf{C^*\text{-}Alg}$: the equalizer of $\alpha, \beta : \mathcal{A} \to \mathcal{B}$ can be described as the inclusion map in $\mathcal{A}$ of the maximal C^*-subalgebra of $\mathcal{A}$ on which α and β agree; given a family $(\alpha_i)_{i\in I}$ of morphisms $\alpha_i : \mathcal{A}_i \to \mathcal{B}$, their intersection $\bigwedge_{i\in I} \alpha_i$ is the inclusion map of the set-theoretic intersection $\bigcap_{i\in I} \alpha_i(\mathcal{A}_i)$ in $\mathcal{B}$, while the union $\bigvee_{i\in I} \alpha_i$ is the inclusion of the C^*-subalgebra of $\mathcal{B}$ generated by the $\alpha_i(\mathcal{A}_i)$ [i.e., the intersection of all C^*-subalgebras containing the set-theoretic union].

3.2 The category of locally covariant theories

Once the category Phys has been selected, we may follow the line of BFV and define a locally covariant physical theory of the given type to be any (covariant) functor $\mathscr{A}$ from Loc to Phys (equally we may use Loc_0 as the domain category if we wish to restrict to connected spacetimes). Thus, to each spacetime $\boldsymbol{M} \in \mathsf{Loc}$ there is an object $\mathscr{A}(\boldsymbol{M})$ of Phys and to each Loc-morphism $\psi : \boldsymbol{M} \to \boldsymbol{N}$ there is an Phys-morphism $\mathscr{A}(\psi) : \mathscr{A}(\boldsymbol{M}) \to \mathscr{A}(\boldsymbol{N})$ such that $\mathscr{A}(\psi \circ \psi') = \mathscr{A}(\psi) \circ \mathscr{A}(\psi')$ for arbitrary compositions of morphisms and $\mathscr{A}(\mathrm{id}_{\boldsymbol{M}}) = \mathrm{id}_{\mathscr{A}(\boldsymbol{M})}$ for all $\boldsymbol{M}$. For BFV, where Phys is a suitable category of $*$-algebras, the $\mathscr{A}(\boldsymbol{M})$ is the algebra of observables or of smeared fields describing the theory in spacetime $\boldsymbol{M}$.

There is always at least one theory, namely the trivial theory $\mathscr{I}$ with $\mathscr{I}(\boldsymbol{M}) = \mathcal{I}$, $\mathscr{I}(\psi) = \mathrm{id}_{\mathcal{I}}$, where $\mathcal{I}$ is the initial object of Phys. As shown in BFV, the standard example of the Klein–Gordon field provides another example (with Phys chosen as a category of $*$- or C^*-algebras according to the quantization method); the same is true of the extended algebra of Wick products [32] (refined, to remove the dependence on choice of Hadamard function, as in [8, §5.5.3]) and (passing to the category of globally hyperbolic spacetimes with spin structure) the Dirac field [48] and its corresponding extended algebra [19]. (Strictly, these examples were discussed in the context of functors from Loc_0 to Alg or $\mathsf{C^*\text{-}Alg}$, but they generalise to Loc.)

The functorial nature of a theory $\mathscr{A}$ ensures that it respects local general covariance, as we will see in Sect. 3.3. In practice various other properties would normally be expected of the theory. Here, the most important will be the *time-slice property* which requires that $\mathscr{A}$ maps Cauchy morphisms of Loc to isomorphisms in Alg.[12] The time-slice property

[12] In BFV, the timeslice property was phrased in terms of surjectivity of $\mathscr{A}(\psi)$ – an equivalent formulation in the category of C^*-algebras. In general, however, what is needed is the invertibility of $\mathscr{A}(\psi)$ (in Phys) when ψ is Cauchy.

essentially asserts the existence of a dynamical law for the theory and will hold in this form for many different physical theories.

For our purposes, it will be important to regard locally covariant theories as objects within the functor category $\mathsf{LCT} = \mathrm{Funct}(\mathsf{Loc}, \mathsf{Phys})$ (or, $\mathsf{LCT}_0 = \mathrm{Funct}(\mathsf{Loc}_0, \mathsf{Phys})$) in which the morphisms are natural transformations $\zeta : \mathscr{A} \stackrel{.}{\to} \mathscr{B}$. Thus, to each $\boldsymbol{M} \in \mathsf{Loc}$ there is a morphism $\zeta_{\boldsymbol{M}} : \mathscr{A}(\boldsymbol{M}) \to \mathscr{B}(\boldsymbol{M})$ such that $\mathscr{B}(\psi) \circ \zeta_{\boldsymbol{M}} = \zeta_{\boldsymbol{N}} \circ \mathscr{A}(\psi)$ for all morphisms $\psi : \boldsymbol{M} \to \boldsymbol{N}$. The physical interpretation of a morphism $\zeta : \mathscr{A} \stackrel{.}{\to} \mathscr{B}$ is that it provides a way of embedding the theory $\mathscr{A}$ as a subtheory of $\mathscr{B}$. In the special case where every component $\zeta_{\boldsymbol{M}}$ is an isomorphism, ζ is said to be a natural isomorphism; we interpret this as indicating that the theories are equivalent.

Simple examples of morphisms in LCT may be constructed as follows. First, the trivial theory $\mathscr{I}$ is a subtheory of every theory $\mathscr{A}$, and indeed is an initial object for LCT because there is a unique natural $\mathscr{I}_{\mathscr{A}} : \mathscr{I} \stackrel{.}{\to} \mathscr{A}$, whose typical component is $(\mathscr{I}_{\mathscr{A}})_{\boldsymbol{M}} = \mathcal{I}_{\mathscr{A}(\boldsymbol{M})}$, the unique morphism $\mathcal{I} \to \mathscr{A}(\boldsymbol{M})$.

Second, given an endofunctor $\mathscr{F}$ of Phys and a natural $\eta : \mathscr{F} \stackrel{.}{\to} \mathrm{id}_{\mathsf{Phys}}$, any $\mathscr{A} \in \mathsf{LCT}$ has a subtheory $\zeta : \mathscr{F} \circ \mathscr{A} \stackrel{.}{\to} \mathscr{A}$. In the case $\mathsf{Phys} = \mathsf{TAlg}$, an example is given as follows: to each object $\mathcal{A}$, let $\mathscr{F}(\mathcal{A})$ be the same $*$-algebra but equipped with the discrete topology, and let $\eta_{\mathcal{A}} : \mathscr{F}(\mathcal{A}) \to \mathcal{A}$ have the identity as its underlying $*$-homomorphism; then $\mathscr{F}$ can be defined on morphisms in the obvious way so as to ensure naturality of $\eta : \mathscr{F} \stackrel{.}{\to} \mathrm{id}_{\mathsf{TAlg}}$.

Third, if Phys is a monoidal category (see, e.g., [41]), with the initial object as the unit, it induces a monoidal structure on LCT: given $\mathscr{A}, \mathscr{B} \in \mathsf{LCT}$, define $(\mathscr{A} \otimes \mathscr{B})(\boldsymbol{M}) = \mathscr{A}(\boldsymbol{M}) \otimes \mathscr{B}(\boldsymbol{M})$ and $(\mathscr{A} \otimes \mathscr{B})(\psi) = \mathscr{A}(\psi) \otimes \mathscr{B}(\psi)$; this is easily checked to define a new functor $\mathscr{A} \otimes \mathscr{B} \in \mathsf{LCT}$. The theory $\mathscr{I}$ is the unit for the tensor product in LCT and the associators and unitors all lift immediately. For example, recall that the right unitor ρ of a monoidal category Phys is a natural isomorphism with components $\rho_{\mathcal{A}} : \mathcal{A} \otimes \mathcal{I} \cong \mathcal{A}$ (obeying certain properties). This lifts to a natural $\hat{\rho}$, with components $\hat{\rho}_{\mathscr{A}} : \mathscr{A} \otimes \mathscr{I} \cong \mathscr{A}$, where $(\hat{\rho}_{\mathscr{A}})_{\boldsymbol{M}} = \rho_{\mathscr{A}(\boldsymbol{M})}$ and which functions as the right unitor in LCT. One may check that all the coherence properties required of a monoidal structure lift in this way. Writing $\hat{\lambda}$ for the left-unitor in LCT, we obtain LCT morphisms $\eta_{\mathscr{A},\mathscr{B}} : \mathscr{A} \stackrel{.}{\to} \mathscr{A} \otimes \mathscr{B}$ and $\zeta_{\mathscr{A},\mathscr{B}} : \mathscr{B} \stackrel{.}{\to} \mathscr{A} \otimes \mathscr{B}$ for any pair of theories $\mathscr{A}, \mathscr{B} \in \mathsf{LCT}$, given by

$$\eta_{\mathscr{A},\mathscr{B}} = (\mathrm{id}_{\mathscr{A}} \otimes \mathscr{I}_{\mathscr{B}}) \circ \hat{\rho}_{\mathscr{A}}^{-1} \qquad \zeta_{\mathscr{A},\mathscr{B}} = (\mathscr{I}_{\mathscr{A}} \otimes \mathrm{id}_{\mathscr{B}}) \circ \hat{\lambda}_{\mathscr{B}}^{-1}.$$

Given these structures we can define arbitrary monoidal powers of a given theory $\mathscr{A} \in \mathsf{LCT}$, by setting, for example, $\mathscr{A}^{\otimes 1} := \mathscr{A}$ and $\mathscr{A}^{\otimes (k+1)} := \mathscr{A}^{\otimes k} \otimes \mathscr{A}$ for each $k \in \mathbb{N}$.[13] Then $\gamma(k) := \eta_{\mathscr{A}^{\otimes k},\mathscr{A}}$ provides a natural transformation $\gamma(k) : \mathscr{A}^{\otimes k} \stackrel{.}{\to} \mathscr{A}^{\otimes (k+1)}$; and if $k < k'$ are any natural numbers we may set

$$\beta(k, k') = \gamma(k' - 1) \circ \cdots \circ \gamma(k), \tag{3.1}$$

giving a natural transformation $\beta(k, k') : \mathscr{A}^{\otimes k} \stackrel{.}{\to} \mathscr{A}^{\otimes k'}$. Defining, additionally, $\beta(k, k) = \mathrm{id}_{\mathscr{A}^{\otimes k}}$ (the identity morphism of $\mathscr{A}^{\otimes k}$ in LCT), it is clear that $\beta(k', k'') \circ \beta(k, k') = \beta(k, k'')$

[13] Thus $\mathscr{A}^{\otimes k} = ((\cdots((\mathscr{A} \otimes \mathscr{A}) \otimes \mathscr{A}) \otimes \cdots) \otimes \mathscr{A}) \otimes \mathscr{A}$. In a monoidal category in which associators are not necessarily identities, there would be other possible (isomorphic) definitions of the monoidal powers by different placement of brackets.

whenever $k \leq k' \leq k''$ and that β defines a functor $\beta : \mathsf{N} \to \mathsf{LCT}$. Here, N is the category whose morphisms are ordered pairs (k, k') of natural numbers with $k \leq k'$ and composition $(k', k'') \circ (k, k') = (k, k'')$; that is, N is the partially ordered set $(\mathbb{N}, \leq)$ regarded as a category.

In the case $\mathsf{Phys} = \mathsf{Alg}$, using the algebraic tensor product, these constructions reduce to

$$\mathscr{A}^{\otimes k}(\boldsymbol{M}) = \mathscr{A}(\boldsymbol{M})^{\otimes k}, \qquad \beta(k,k')_{\boldsymbol{M}} A = A \otimes \mathbf{1}_{\mathscr{A}(\boldsymbol{M})}^{\otimes(k'-k)} \quad (A \in \mathscr{A}(\boldsymbol{M})^{\otimes k})$$

for $k < k'$; the duals of the $\beta(k, k')_{\boldsymbol{M}}$ are of course partial traces.

There are many similar ways of constructing functors from N to LCT, of course, but the above will suffice for our purposes and provide useful building blocks in the sequel.

3.3 The kinematic net

One of the aims of the BFV paper was to formulate QFT in curved spacetime in such a way that algebraic quantum field theory in Minkowski space could be recovered as a special case. This requires that every suitable subregion of a spacetime $\boldsymbol{M}$ should be associated with a subalgebra of the algebra $\mathscr{A}(\boldsymbol{M})$ assigned to $\boldsymbol{M}$ by the theory $\mathscr{A}$ (for the moment, we take $\mathsf{Phys} = \mathsf{Alg}$ or $\mathsf{C^*\text{-}Alg}$).

For any $\boldsymbol{M} \in \mathsf{Loc}$, recall that $\mathscr{O}(\boldsymbol{M})$ is the set of globally hyperbolic open subsets of $\boldsymbol{M}$ with at most finitely many connected components, all of which are mutually causally disjoint, and $\mathscr{O}_0(\boldsymbol{M})$ those which are connected. For each nonempty $O \in \mathscr{O}(\boldsymbol{M})$ (resp., $O \in \mathscr{O}_0(\boldsymbol{M})$), we have a canonical inclusion $\iota_{\boldsymbol{M};O} : \boldsymbol{M}|_O \to \boldsymbol{M}$, an algebra $\mathscr{A}(\boldsymbol{M}|_O)$ and a morphism $\mathscr{A}(\iota_{\boldsymbol{M};O}) : \mathscr{A}(\boldsymbol{M}|_O) \to \mathscr{A}(\boldsymbol{M})$. BFV took the image of $\mathscr{A}(\iota_{\boldsymbol{M};O})$ as the subalgebra associated with O and showed that this assignment generalises AQFT. To facilitate the discussion of arbitrary categories Phys it is better to focus attention on the morphism $\mathscr{A}(\iota_{\boldsymbol{M};O})$ than its 'image' (which is not defined in general categories).

Accordingly, let Phys be any category obeying our minimal assumptions and let $\mathscr{A} \in \mathsf{LCT}$ (resp., LCT_0). For $\boldsymbol{M} \in \mathsf{Loc}$ (resp., Loc_0) and nonempty $O \in \mathscr{O}(\boldsymbol{M})$ (resp., $O \in \mathscr{O}_0(\boldsymbol{M})$), we define

$$\mathscr{A}^{\mathrm{kin}}(\boldsymbol{M}; O) = \mathscr{A}(\boldsymbol{M}|_O), \qquad \text{and} \qquad \alpha^{\mathrm{kin}}_{\boldsymbol{M};O} = \mathscr{A}(\iota_{\boldsymbol{M};O}) : \mathscr{A}^{\mathrm{kin}}(\boldsymbol{M}; O) \to \mathscr{A}(\boldsymbol{M}).$$

We refer to the assignment $O \mapsto \alpha^{\mathrm{kin}}_{\boldsymbol{M};O}$ as the *kinematic net*. Strictly, BFV only considered local algebras corresponding to relatively compact globally hyperbolic subsets; however it is useful (and natural, in the functorial setting) to extend the assignment of local algebras to regions with noncompact closure. Note, however, that the pathologies discussed below are already visible for local algebras of relatively compact regions.

The following result shows that the subobject depends only on $\boldsymbol{M}$ and O. In the statement of this result, the $\cong$ symbol between two morphisms with a common codomain asserts their isomorphism as subobjects of the codomain object; i.e., $\alpha \cong \beta$ holds iff there is a (necessarily unique) isomorphism γ such that $\alpha = \beta \circ \gamma$; see Appendix B.

Lemma 3.1 *If $\psi : \boldsymbol{L} \to \boldsymbol{M}$ then $\mathscr{A}(\psi) \cong \alpha^{\mathrm{kin}}_{\boldsymbol{M};\psi(\boldsymbol{L})}$.*

Proof: We may factor $\psi = \iota_{\boldsymbol{M};\psi(\boldsymbol{L})} \circ \tilde{\psi}$ where $\tilde{\psi} : \boldsymbol{L} \to \boldsymbol{M}|_{\psi(\boldsymbol{L})}$ is an isomorphism; as functors preserve isomorphisms we therefore have $\mathscr{A}(\psi) = \alpha^{\rm kin}_{\boldsymbol{M};\psi(\boldsymbol{L})} \circ \mathscr{A}(\tilde{\psi}) \cong \alpha^{\rm kin}_{\boldsymbol{M};\psi(\boldsymbol{L})}$. $\square$

The basic properties of $O \mapsto \mathscr{A}^{\rm kin}(\boldsymbol{M};O)$ are discussed in Prop. 2.3 of BFV [in the case of connected O]. In particular, if $O_1 \subset O_2$ then $\iota_{\boldsymbol{M};O_1}$ factorises via $\iota_{\boldsymbol{M};O_2}$ as $\iota_{\boldsymbol{M};O_1} = \iota_{\boldsymbol{M};O_2} \circ \iota_{\boldsymbol{M}|_{O_2};O_1}$ and the functorial property of $\mathscr{A}$ implies

$$\alpha^{\rm kin}_{\boldsymbol{M};O_1} = \alpha^{\rm kin}_{\boldsymbol{M};O_2} \circ \mathscr{A}(\iota_{\boldsymbol{M}|_{O_2};O_1}), \tag{3.2}$$

which can also be written in the form $\alpha^{\rm kin}_{\boldsymbol{M};O_1} \le \alpha^{\rm kin}_{\boldsymbol{M};O_2}$, where $\le$ is the order relation in the subobject lattice of $\mathscr{A}(\boldsymbol{M})$ (see, e.g., [21]). That is, the kinematic net is *isotonous*.

If, additionally, O_1 contains a Cauchy surface for O_2, then the morphism $\iota_{\boldsymbol{M}|_{O_2};O_1}$ is Cauchy and is mapped to an isomorphism if $\mathscr{A}$ obeys the timeslice property. Then the factorisation (3.2) asserts that $\alpha^{\rm kin}_{\boldsymbol{M};O_1}$ and $\alpha^{\rm kin}_{\boldsymbol{M};O_2}$ determine isomorphic subobjects of $\mathscr{A}(\boldsymbol{M})$: we write $\alpha^{\rm kin}_{\boldsymbol{M};O_1} \cong \alpha^{\rm kin}_{\boldsymbol{M};O_2}$. [This is an improved formulation of Prop. 2.3(d) in BFV. Compare also Thm. 5.4 below.]

Now suppose that $\psi : \boldsymbol{M} \to \boldsymbol{N}$. If $O \in \mathscr{O}(\boldsymbol{M})$ (resp., $\mathscr{O}_0(\boldsymbol{M})$) is nonempty then $\psi(O) \in \mathscr{O}(\boldsymbol{N})$ (resp., $\mathscr{O}_0(\boldsymbol{N})$) and there is an isomorphism $\tilde{\psi} : \boldsymbol{M}|_O \to \boldsymbol{N}|_{\psi(O)}$ such that $\psi \circ \iota_{\boldsymbol{M};O} = \iota_{\boldsymbol{N};\psi(O)} \circ \tilde{\psi}$. Applying the functor $\mathscr{A}$, this gives a commuting diagram

$$\begin{tikzcd}
\mathscr{A}^{\rm kin}(\boldsymbol{M};O) \arrow[r, "\mathscr{A}(\tilde{\psi})", "\cong"'] \arrow[d, "\alpha^{\rm kin}_{\boldsymbol{M};O}"'] & \mathscr{A}^{\rm kin}(\boldsymbol{N};\psi(O)) \arrow[d, "\alpha^{\rm kin}_{\boldsymbol{N};\psi(O)}"] \\
\mathscr{A}(\boldsymbol{M}) \arrow[r, "\mathscr{A}(\psi)"] & \mathscr{A}(\boldsymbol{N})
\end{tikzcd}$$

and thus the equivalence of subobjects

$$\alpha^{\rm kin}_{\boldsymbol{N};\psi(O)} \cong \mathscr{A}(\psi) \circ \alpha^{\rm kin}_{\boldsymbol{M};O}, \tag{3.3}$$

which expresses the covariance of the kinematic net. In particular, this gives the action of automorphisms of $\boldsymbol{M}$ (i.e., a (time-)orientation preserving isometric diffeomorphism) on the kinematic net: the functor $\mathscr{A}$ provides a representation of the automorphism group $\mathrm{Aut}(\boldsymbol{M})$ in the automorphism group of $\mathscr{A}(\boldsymbol{M})$ by $\kappa \mapsto \mathscr{A}(\kappa)$, and the formula

$$\alpha^{\rm kin}_{\boldsymbol{M};\kappa(O)} \cong \mathscr{A}(\kappa) \circ \alpha^{\rm kin}_{\boldsymbol{M};O}$$

shows that this has the expected geometrical action on the kinematic net.

3.4 Relative Cauchy evolution

Let $\boldsymbol{M} = (\mathcal{M}, \boldsymbol{g}, \mathfrak{o}, \mathfrak{t}) \in \mathsf{Loc}$ be a globally hyperbolic spacetime. Given any symmetric $\boldsymbol{h} \in C_0^\infty(T^0_2\boldsymbol{M})$ such that $\boldsymbol{g} + \boldsymbol{h}$ is a time-orientable Lorentz metric on $\mathcal{M}$, there is a unique choice of time-orientation $\mathfrak{t}_{\boldsymbol{h}}$ for $\boldsymbol{g} + \boldsymbol{h}$ that agrees with $\mathfrak{t}$ outside K. If $\boldsymbol{M}[\boldsymbol{h}] =$

$(\mathcal{M}, \boldsymbol{g}+\boldsymbol{h}, \mathfrak{o}, \mathfrak{t}_{\boldsymbol{h}})$ is a globally hyperbolic spacetime, we say that $\boldsymbol{h}$ is a *globally hyperbolic perturbation* of $\boldsymbol{M}$ and write $\boldsymbol{h} \in H(\boldsymbol{M})$. The subset of $\boldsymbol{h} \in H(\boldsymbol{M})$ with support in $K \subset \mathcal{M}$ is denoted $H(\boldsymbol{M}; K)$. Clearly, $\boldsymbol{M} = \boldsymbol{M}[\mathbf{0}]$, where $\mathbf{0}$ is identically zero, so $H(\boldsymbol{M})$ is nonempty; in fact it contains an open neighbourhood of $\mathbf{0}$ in the usual test-function topology on symmetric smooth compactly supported sections of $T^0_2\boldsymbol{M}$ (see §7.1 of [2]). We endow $H(\boldsymbol{M})$ with the subspace topology induced from $\mathscr{D}(T^0_2 M)$.

If a theory $\mathscr{A} \in \mathsf{LCT}$ has the timeslice property then, as shown by BFV, we may compare the dynamics on $\boldsymbol{M}$ and its perturbations via a *relative Cauchy evolution*. We now describe the construction in more depth than BFV, paying attention to the covariance properties of the relative Cauchy evolution and the relation between the evolutions of theories related by morphisms in LCT. A number of geometrical lemmas, including the following, will be proved in Appendix A.

Lemma 3.2 *Let $\boldsymbol{M} \in \mathsf{Loc}$ and $\boldsymbol{h} \in H(\boldsymbol{M})$, and set $\mathcal{M}^\pm = \mathcal{M} \setminus J^\mp_{\boldsymbol{M}}(\operatorname{supp}\boldsymbol{h})$. Then (a) $\mathcal{M}^\pm$ are globally hyperbolic subsets of both $\boldsymbol{M}$ and $\boldsymbol{M}[\boldsymbol{h}]$, and $\boldsymbol{M}^\pm[\boldsymbol{h}] \stackrel{\text{def}}{=} \boldsymbol{M}|_{\mathcal{M}^\pm} = \boldsymbol{M}[\boldsymbol{h}]|_{\mathcal{M}^\pm}$; (b) the canonical inclusions $\imath^\pm_{\boldsymbol{M}}[\boldsymbol{h}] \stackrel{\text{def}}{=} \iota_{\boldsymbol{M};\mathcal{M}^\pm}$ and $\jmath^\pm_{\boldsymbol{M}}[\boldsymbol{h}] \stackrel{\text{def}}{=} \iota_{\boldsymbol{M}[\boldsymbol{h}];\mathcal{M}^\pm}$ are Cauchy morphisms. If $\boldsymbol{M} \in \mathsf{Loc}_0$, then $\boldsymbol{M}[\boldsymbol{h}]$, $\boldsymbol{M}^\pm[\boldsymbol{h}]$ are also in Loc_0, and the morphisms $\iota_{\boldsymbol{M}}[\boldsymbol{h}]^\pm$, $\jmath^\pm_{\boldsymbol{M}}[\boldsymbol{h}]$ are Loc_0-morphisms.*

Among other things, this result shows that we can work consistently in either Loc or Loc_0. For the rest of this section, we will not distinguish between Loc or Loc_0 in the statement of our results (with the exception of Prop. 3.5, where there is a slight difference) but it should be understood that all spacetimes and morphisms should be taken consistently from one or other of Loc or Loc_0, and that the locally covariant theories mentioned are taken consistently from LCT or LCT_0 respectively. (In some cases, the proofs of these statements differ slightly depending on which category is being used.)

Proceeding in this way, if $\boldsymbol{M}$ is a spacetime, each $\boldsymbol{h} \in H(\boldsymbol{M})$ induces a *past Cauchy wedge*, i.e., the diagram

$$\boldsymbol{M} \xleftarrow{\imath^-_{\boldsymbol{M}}[\boldsymbol{h}]} \boldsymbol{M}^-[\boldsymbol{h}] \xrightarrow{\jmath^-_{\boldsymbol{M}}[\boldsymbol{h}]} \boldsymbol{M}[\boldsymbol{h}]$$

and a *future Cauchy wedge*, namely,

$$\boldsymbol{M} \xleftarrow{\imath^+_{\boldsymbol{M}}[\boldsymbol{h}]} \boldsymbol{M}^+[\boldsymbol{h}] \xrightarrow{\jmath^+_{\boldsymbol{M}}[\boldsymbol{h}]} \boldsymbol{M}[\boldsymbol{h}].$$

Any locally covariant theory $\mathscr{A}$ obeying the timeslice axiom will map each morphism in the past and future Cauchy wedges to an isomorphism. In particular there are isomorphisms

$$\tau^\pm_{\boldsymbol{M}}[\boldsymbol{h}] = \mathscr{A}(\jmath^\pm_{\boldsymbol{M}}[\boldsymbol{h}]) \circ (\mathscr{A}(\imath^\pm_{\boldsymbol{M}}[\boldsymbol{h}]))^{-1} : \mathscr{A}(\boldsymbol{M}) \to \mathscr{A}(\boldsymbol{M}[\boldsymbol{h}])$$

and an automorphism $\mathrm{rce}_{\boldsymbol{M}}[\boldsymbol{h}]$ of $\mathscr{A}(\boldsymbol{M})$ given by

$$\mathrm{rce}_{\boldsymbol{M}}[\boldsymbol{h}] = (\tau^-_{\boldsymbol{M}}[\boldsymbol{h}])^{-1} \circ \tau^+_{\boldsymbol{M}}[\boldsymbol{h}],$$

which is called the *relative Cauchy evolution* induced by $\boldsymbol{h}$. Not all metric perturbations are physically significant: for example, if $\boldsymbol{M} = (\mathcal{M}, \boldsymbol{g}, \mathfrak{o}, \mathfrak{t})$ and ψ is a diffeomorphism of $\mathcal{M}$

acting as the identity outside a compact set K, then ψ induces a morphism (also denoted ψ) from $\boldsymbol{M}$ to $\boldsymbol{M}' = (\mathcal{M}, \psi_*\boldsymbol{g}, \psi_*\mathfrak{o}, \psi_*\mathfrak{t})$ which can be regarded as a globally hyperbolic perturbation $\boldsymbol{M}' = \boldsymbol{M}[\boldsymbol{h}]$ for $\boldsymbol{h} = \psi_*\boldsymbol{g} - \boldsymbol{g}$. It is easily seen that

$$\psi \circ \imath^{\pm}_{\boldsymbol{M}}[\boldsymbol{h}] = \jmath^{\pm}_{\boldsymbol{M}}[\boldsymbol{h}]$$

for both choices of sign; accordingly, we have $\tau^{+}_{\boldsymbol{M}}[\boldsymbol{h}] = \mathscr{A}(\psi) = \tau^{-}_{\boldsymbol{M}}[\boldsymbol{h}]$ and hence $\mathrm{rce}_{\boldsymbol{M}}[\boldsymbol{h}] = \mathrm{id}_{\mathscr{A}(\boldsymbol{M})}$, which reflects the fact that $\boldsymbol{M}$ and $\boldsymbol{M}'$ are physically equivalent and have equivalent dynamics.

The definition of relative Cauchy evolution given here differs slightly from that given in BFV, where the Cauchy morphisms used were not fixed by the perturbation; our approach avoids the necessity of demonstrating that the definition does not depend on the choices made by introducing the preferred past and future Cauchy wedges. In order to make contact with the original definition, however, we give the following result, which is also useful for computations (and, in passing, establishes the independence mentioned above).

Proposition 3.3 *Let K be a compact subset of $\mathcal{M}$ and suppose $\psi^{\pm} : \boldsymbol{L}^{\pm} \to \boldsymbol{M}$ are Cauchy morphisms with image contained in $\mathcal{M} \setminus J^{\mp}_{\boldsymbol{M}}(K)$. For each $\boldsymbol{h} \in H(\boldsymbol{M}; K)$ there are morphisms $\psi^{\pm}[\boldsymbol{h}] : \boldsymbol{L}^{\pm} \to \boldsymbol{M}[\boldsymbol{h}]$ with the same underlying embedding as $\psi^{\pm}$ such that*

$$\tau^{\pm}_{\boldsymbol{M}}[\boldsymbol{h}] = \mathscr{A}(\psi^{\pm}[\boldsymbol{h}]) \circ \mathscr{A}(\psi^{\pm})^{-1}$$

and hence

$$\mathrm{rce}_{\boldsymbol{M}}[\boldsymbol{h}] = \mathscr{A}(\psi^{-}) \circ \mathscr{A}(\psi[\boldsymbol{h}]^{-})^{-1} \circ \mathscr{A}(\psi^{+}[\boldsymbol{h}]) \circ \mathscr{A}(\psi^{+})^{-1}.$$

Proof: The following lemma will be proved in Appendix A.

Lemma 3.4 *Let K be a compact subset of $\mathcal{M}$ and suppose $\psi : \boldsymbol{L} \to \boldsymbol{M}$ has its range contained in one or both of $\mathcal{M} \setminus J^{\mp}_{\boldsymbol{M}}(K)$. Then the underlying embedding of ψ induces $\psi[\boldsymbol{h}] : \boldsymbol{L} \to \boldsymbol{M}[\boldsymbol{h}]$ for any $\boldsymbol{h} \in H(\boldsymbol{M}; K)$ (independent of the choice $\pm$ in the hypothesis). Moreover, there is $\varphi^{\pm} : \boldsymbol{L} \to \boldsymbol{M}^{\pm}[\boldsymbol{h}]$ such that*

$$\psi = \imath^{\pm}_{\boldsymbol{M}}[\boldsymbol{h}] \circ \varphi^{\pm}, \qquad \psi[\boldsymbol{h}] = \jmath^{\pm}_{\boldsymbol{M}}[\boldsymbol{h}] \circ \varphi^{\pm}.$$

If ψ is Cauchy then so are $\psi[\boldsymbol{h}]$ and $\varphi^{\pm}$.

The immediate consequence is that

$$\tau^{\pm}_{\boldsymbol{M}}[\boldsymbol{h}] \circ \mathscr{A}(\psi) = \tau^{\pm}_{\boldsymbol{M}}[\boldsymbol{h}] \circ \mathscr{A}(\imath^{\pm}_{\boldsymbol{M}}[\boldsymbol{h}]) \circ \mathscr{A}(\varphi^{\pm}) = \mathscr{A}(\jmath^{\pm}_{\boldsymbol{M}}[\boldsymbol{h}]) \circ \mathscr{A}(\varphi^{\pm}) = \mathscr{A}(\psi[\boldsymbol{h}]).$$

Applying the $+$ (resp., $-$) case to the ψ^{+} (resp., ψ^{-}) in the hypothesis, Prop. 3.3 follows. $\square$

Much of the present paper depends crucially on locality and covariance properties of the relative Cauchy evolution that were not addressed in BFV. Locality can be obtained from Lem. 3.4.

Proposition 3.5 *Let K be a compact subset of $\mathcal{M}$ and suppose $\psi : \boldsymbol{L} \to \boldsymbol{M}$ has its range contained in the causal complement $K^\perp = \mathcal{M} \setminus J_{\boldsymbol{M}}(K)$ of K (hence, in particular, if $K \subset \psi(\boldsymbol{L})'$). Then*

$$\mathrm{rce}_{\boldsymbol{M}}[\boldsymbol{h}] \circ \mathscr{A}(\psi) = \mathscr{A}(\psi)$$

for all $\boldsymbol{h} \in H(\boldsymbol{M}; K)$. In particular, this implies that the kinematical net obeys

$$\mathrm{rce}_{\boldsymbol{M}}[\boldsymbol{h}] \circ \alpha^{\mathrm{kin}}_{\boldsymbol{M};O} = \alpha^{\mathrm{kin}}_{\boldsymbol{M};O}$$

for all nonempty $O \in \mathscr{O}(\boldsymbol{M})$ (or $\mathscr{O}_0(\boldsymbol{M})$ for theories in LCT_0) with $O \subset (\mathrm{supp}\,\boldsymbol{h})^\perp$.

Proof: The morphism ψ obeys the hypothesis of Lem. 3.4 in both the $+$ and $-$ cases. Accordingly

$$\tau^+_{\boldsymbol{M}}[\boldsymbol{h}] \circ \mathscr{A}(\psi) = \mathscr{A}(\psi[\boldsymbol{h}]) = \tau^-_{\boldsymbol{M}}[\boldsymbol{h}] \circ \mathscr{A}(\psi)$$

and the result follows on composing with $\tau^-_{\boldsymbol{M}}[\boldsymbol{h}]^{-1}$. □

We remark that the hypotheses of this result allow for nontrivial intersection of $\mathrm{cl}(\psi(\boldsymbol{L}))$ and $J_{\boldsymbol{M}}(K)$.

Our covariance result depends on the following geometrical lemma.

Lemma 3.6 *For each morphism $\psi : \boldsymbol{M} \to \boldsymbol{N}$, we have $\psi_* H(\boldsymbol{M}) \subset H(\boldsymbol{N})$. Moreover, for each $\boldsymbol{h} \in H(\boldsymbol{M})$ and $\psi : \boldsymbol{M} \to \boldsymbol{N}$ there are morphisms $\psi^\pm[\boldsymbol{h}] : \boldsymbol{M}^\pm[\boldsymbol{h}] \to \boldsymbol{N}^\pm[\psi_*\boldsymbol{h}]$ and $\psi[\boldsymbol{h}] : \boldsymbol{M}[\boldsymbol{h}] \to \boldsymbol{N}[\psi_*\boldsymbol{h}]$ so that the following diagram commutes:*

$$\begin{tikzcd} \boldsymbol{M} \arrow[d, "\psi"] & \boldsymbol{M}^\pm[\boldsymbol{h}] \arrow[l, "\imath^\pm_{\boldsymbol{M}}[\boldsymbol{h}]"'] \arrow[r, "\jmath^\pm_{\boldsymbol{M}}[\boldsymbol{h}]"] \arrow[d, "\psi^\pm[\boldsymbol{h}]"] & \boldsymbol{M}[\boldsymbol{h}] \arrow[d, "\psi[\boldsymbol{h}]"] \\ \boldsymbol{N} & \boldsymbol{N}^\pm[\psi_*\boldsymbol{h}] \arrow[l, "\imath^\pm_{\boldsymbol{N}}[\psi_*\boldsymbol{h}]"] \arrow[r, "\jmath^\pm_{\boldsymbol{N}}[\psi_*\boldsymbol{h}]"'] & \boldsymbol{N}[\psi_*\boldsymbol{h}] \end{tikzcd} . \tag{3.4}$$

Proof: The most involved aspect is to prove that $\psi_*\boldsymbol{h} \in H(\boldsymbol{N})$. This is accomplished by Lem. A.7 below. As the horizontal morphisms in diagram (3.4) are inclusions it is now sufficient to show that there are morphisms $\psi^\pm[\boldsymbol{h}] : \boldsymbol{M}^\pm[\boldsymbol{h}] \to \boldsymbol{N}^\pm[\psi_*\boldsymbol{h}]$ and $\psi[\boldsymbol{h}] : \boldsymbol{M}[\boldsymbol{h}] \to \boldsymbol{N}[\psi_*\boldsymbol{h}]$ with the same underlying embedding as ψ; the diagram will then automatically commute. The existence of $\psi[\boldsymbol{h}]$ is obvious. As the image of ψ is causally convex in $\boldsymbol{N}$, $J^\mp_{\boldsymbol{N}}(\mathrm{supp}\,\psi_*\boldsymbol{h}) \cap \psi(\mathcal{M}) = \psi(J^\mp_{\boldsymbol{M}}(\mathrm{supp}\,\boldsymbol{h}))$ and hence $\psi(\mathcal{M} \setminus J^\mp_{\boldsymbol{M}}(\mathrm{supp}\,\boldsymbol{h})) \subset \mathcal{N} \setminus J^\mp_{\boldsymbol{N}}(\mathrm{supp}\,\psi_*\boldsymbol{h}))$. Hence the underlying embedding induces $\psi^\pm[\boldsymbol{h}] : \boldsymbol{M}^\pm[\boldsymbol{h}] \to \boldsymbol{N}^\pm[\boldsymbol{h}]$ as required. □

This result shows that the sets of hyperbolic perturbations are functorially assigned to spacetimes of Loc and Loc_0, and the push-forward induces a mapping between Cauchy wedges, which could also be interpreted as a morphism in a suitable category of wedges. We do not pursue this here. The main use of the above lemma is to establish covariance of the relative Cauchy evolution.

Proposition 3.7 *If $\psi : \boldsymbol{M} \to \boldsymbol{N}$ and $\boldsymbol{h} \in H(\boldsymbol{M})$, then*

$$\tau^{\pm}_{\boldsymbol{N}}[\psi_*\boldsymbol{h}] \circ \mathscr{A}(\psi) = \mathscr{A}(\psi[\boldsymbol{h}]) \circ \tau^{\pm}_{\boldsymbol{M}}[\boldsymbol{h}], \tag{3.5}$$

and consequently

$$\mathrm{rce}_{\boldsymbol{N}}[\psi_*\boldsymbol{h}] \circ \mathscr{A}(\psi) = \mathscr{A}(\psi) \circ \mathrm{rce}_{\boldsymbol{M}}[\boldsymbol{h}]. \tag{3.6}$$

Proof: Lemma 3.6 demonstrates that the $\tau^{\pm}_{\boldsymbol{N}}[\psi_*\boldsymbol{h}]$ and $\mathrm{rce}_{\boldsymbol{N}}[\psi_*\boldsymbol{h}]$ exist. Taking the image under $\mathscr{A}$ of diagram (3.4) and using the definitions of $\tau^{\pm}_{\boldsymbol{M}}[\boldsymbol{h}]$ and $\tau^{\pm}_{\boldsymbol{N}}[\psi_*\boldsymbol{h}]$, we obtain the commutative diagrams

$$\begin{tikzcd}
\mathscr{A}(\boldsymbol{M}) \arrow[rr, "\tau^{\pm}_{\boldsymbol{M}}[\boldsymbol{h}]"] \arrow[ddd, "\mathscr{A}(\psi)"] & & \mathscr{A}(\boldsymbol{M}[\boldsymbol{h}]) \arrow[ddd, "\mathscr{A}(\psi[\boldsymbol{h}])"] \\
& \mathscr{A}(\boldsymbol{M}^{\pm}[\boldsymbol{h}]) \arrow[ul] \arrow[ur] \arrow[d, "\mathscr{A}(\psi^{\pm}[\boldsymbol{h}])"] & \\
& \mathscr{A}(\boldsymbol{N}^{\pm}[\boldsymbol{h}]) \arrow[dl] \arrow[dr] & \\
\mathscr{A}(\boldsymbol{N}) \arrow[rr, "\tau^{\pm}_{\boldsymbol{N}}[\psi_*\boldsymbol{h}]"'] & & \mathscr{A}(\boldsymbol{N}[\boldsymbol{h}])
\end{tikzcd}$$

(we suppress labels on the slanted arrows) from which Eqs. (3.5) and (3.6) follow immediately. □

So far we have defined relative Cauchy evolution for a single theory $\mathscr{A}$ obeying the timeslice property. Where a number of theories are considered, we will distinguish the relative Cauchy evolution and related structures by a superscript to indicate the theory concerned. The relative Cauchy evolution interacts in an elegant way with the morphisms of LCT and LCT_0:

Proposition 3.8 *Suppose locally covariant theories $\mathscr{A}$ and $\mathscr{B}$ both satisfy the timeslice property and let $\zeta : \mathscr{A} \dot{\to} \mathscr{B}$. For any spacetime $\boldsymbol{M}$ and metric perturbation $\boldsymbol{h} \in H(\boldsymbol{M})$ we have*

$$\zeta_{\boldsymbol{M}[\boldsymbol{h}]} \circ \tau^{(\mathscr{A})\pm}_{\boldsymbol{M}}[\boldsymbol{h}] = \tau^{(\mathscr{B})\pm}_{\boldsymbol{M}}[\boldsymbol{h}] \circ \zeta_{\boldsymbol{M}}$$

and therefore

$$\mathrm{rce}^{(\mathscr{B})}_{\boldsymbol{M}}[\boldsymbol{h}] \circ \zeta_{\boldsymbol{M}} = \zeta_{\boldsymbol{M}} \circ \mathrm{rce}^{(\mathscr{A})}_{\boldsymbol{M}}[\boldsymbol{h}].$$

Proof: Introducing the past and future Cauchy wedges as before, we have

$$\begin{aligned}
\tau^{(\mathscr{B})\pm}_{\boldsymbol{M}}[\boldsymbol{h}] \circ \zeta_{\boldsymbol{M}} \circ \mathscr{A}(\imath^{\pm}_{\boldsymbol{M}}[\boldsymbol{h}]) &= \tau^{(\mathscr{B})\pm}_{\boldsymbol{M}}[\boldsymbol{h}] \circ \mathscr{B}(\imath^{\pm}_{\boldsymbol{M}}[\boldsymbol{h}]) \circ \zeta_{\boldsymbol{M}^{\pm}[\boldsymbol{h}]} = \mathscr{B}(\jmath^{\pm}_{\boldsymbol{M}}[\boldsymbol{h}]) \circ \zeta_{\boldsymbol{M}^{\pm}[\boldsymbol{h}]} \\
&= \zeta_{\boldsymbol{M}[\boldsymbol{h}]} \circ \mathscr{A}(\jmath^{\pm}_{\boldsymbol{M}}[\boldsymbol{h}]) = \zeta_{\boldsymbol{M}[\boldsymbol{h}]} \circ \tau^{(\mathscr{A})\pm}_{\boldsymbol{M}}[\boldsymbol{h}] \circ \mathscr{A}(\imath^{\pm}_{\boldsymbol{M}}[\boldsymbol{h}])
\end{aligned}$$

and since $\mathscr{A}(\imath^{\pm}_{\boldsymbol{M}}[\boldsymbol{h}])$ is epic, the first result holds. Hence

$$\tau^{(\mathscr{B})+}_{\boldsymbol{M}}[\boldsymbol{h}] \circ \mathrm{rce}^{(\mathscr{B})}_{\boldsymbol{M}}[\boldsymbol{h}] \circ \zeta_{\boldsymbol{M}} = \tau^{(\mathscr{B})-}_{\boldsymbol{M}}[\boldsymbol{h}] \circ \zeta_{\boldsymbol{M}} = \zeta_{\boldsymbol{M}[\boldsymbol{h}]} \circ \tau^{(\mathscr{A})-}_{\boldsymbol{M}}[\boldsymbol{h}] = \zeta_{\boldsymbol{M}[\boldsymbol{h}]} \circ \tau^{(\mathscr{A})+}_{\boldsymbol{M}}[\boldsymbol{h}] \circ \mathrm{rce}^{(\mathscr{A})}_{\boldsymbol{M}}[\boldsymbol{h}]$$
$$= \tau^{(\mathscr{B})+}_{\boldsymbol{M}}[\boldsymbol{h}] \circ \zeta_{\boldsymbol{M}} \circ \mathrm{rce}^{(\mathscr{A})}_{\boldsymbol{M}}[\boldsymbol{h}]$$

and as $\tau^{(\mathscr{B})+}_{\boldsymbol{M}}[\boldsymbol{h}]$ is monic the second part follows. □

An important observation in BFV is that the functional derivative of the relative Cauchy evolution with respect to the metric can be interpreted as a stress-energy tensor of the theory, so that (in the case $\mathsf{Phys} = \mathsf{Alg}$ or C^*-Alg)

$$[\boldsymbol{T}_{\boldsymbol{M}}[\boldsymbol{f}], A] = 2i\, \frac{d}{ds} \mathrm{rce}_{\boldsymbol{M}}[s\boldsymbol{f}]A \bigg|_{s=0},$$

where $\boldsymbol{T}_{\boldsymbol{M}}$ is the stress-energy tensor in $\boldsymbol{M}$; the left-hand side should be regarded as the definition of a (not necessarily inner) derivation, and we suppress all technicalities regarding the sense in which differentiation is intended. Prop. 3.8 then has an immediate consequence that

$$[\boldsymbol{T}^{(\mathscr{B})}_{\boldsymbol{M}}[\boldsymbol{f}], \zeta_{\boldsymbol{M}} A] = \zeta_{\boldsymbol{M}} [\boldsymbol{T}^{(\mathscr{A})}_{\boldsymbol{M}}[\boldsymbol{f}], A] \tag{3.7}$$

i.e., a subtheory embedding necessarily intertwines the stress-energy tensors of the two theories.

As an immediate application, consider the quantum field theory of the nonminimally coupled scalar field, with field equation $(\Box_{\boldsymbol{M}} + \xi R_{\boldsymbol{M}} + m^2)\varphi = 0$, where $R_{\boldsymbol{M}}$ is the scalar curvature. For each value of the coupling ξ and the mass m, there is a locally covariant theory $\mathscr{A}^{(m,\xi)}$ so that each $\mathscr{A}^{(m,\xi)}(\boldsymbol{M})$ has generators $\Phi^{(m,\xi)}_{\boldsymbol{M}}(f)$ labelled by $f \in C_0^\infty(\boldsymbol{M})$ subject to relations depending only on the field equation and its Green functions (together with basic structures of linearity and complex conjugation). In Ricci-flat spacetimes, therefore, the map $\Phi^{(m,\xi)}_{\boldsymbol{M}}(f) \mapsto \Phi^{(m,\xi')}_{\boldsymbol{M}}(f)$ extends to an isomorphism $\mathscr{A}^{(m,\xi)}(\boldsymbol{M}) \to \mathscr{A}^{(m,\xi')}(\boldsymbol{M})$ for any ξ, ξ'. We shall call this the 'obvious isomorphism'. Similarly, if $\boldsymbol{M}$ has constant scalar curvature, there is an obvious isomorphism $\mathscr{A}^{(m,\xi)}(\boldsymbol{M}) \cong \mathscr{A}^{(m',\xi')}(\boldsymbol{M})$ whenever $m^2 + \xi R_{\boldsymbol{M}} = m'^2 + \xi' R_{\boldsymbol{M}}$. However, none of these isomorphisms (for distinct values of the labels) can be the components of natural transformations between these theories for the simple reason that the commutators of the stress-energy tensor with the smeared fields (which yield further smeared fields) depend nontrivially on the parameters m and ξ *even in spacetimes that have constant or vanishing scalar curvature.* Thus Eq. (3.7) cannot hold if $\zeta_{\boldsymbol{M}}$ is one of these 'obvious' isomorphisms.

4 Failure of SPASs in LCT

The BFV definition of a locally covariant QFT is that it is a functor $\mathscr{A} : \mathsf{Loc} \to \mathsf{Alg}$. However, in the absence of further assumptions this does not fully answer the question

of what it means for the theory to have the same physical content, i.e., to be 'the same theory' in different spacetimes of the same dimension.[14]

A definition of what it means for a single theory to represent the same physics in all spacetimes (abbreviated SPASs) is not easy to give, and risks the introduction of possibly over-restrictive assumptions on the nature of the theory in question. However it seems reasonable that if we are given two theories, each of which represents the same physics in all spacetimes (by some reasonable definition) and these theories coincide in *some* spacetime, then they should coincide in *all* spacetimes. This motivates the following definition, in which we refer to a natural transformation between functors as a *partial isomorphism* if at least one of its components is an isomorphism.

Definition 4.1 *A class of theories $\mathfrak{T}$ in LCT (or LCT_0) is said to have the* SPASs property *if all partial isomorphisms (in LCT or LCT_0) between theories in $\mathfrak{T}$ are isomorphisms.*

As explained in the introduction, any candidate definition of SPASs can be represented by the class of theories that obey it; the SPASs property can then be used as a necessary criterion on 'good' notions of SPASs. In this section, we will show by examples that neither LCT_0 nor LCT has the SPASs property except where the category Phys is rather trivial; we will use this to demonstrate the existence of individual theories that cannot be regarded as representing the same physics in all spacetimes by any reasonable definition. This will be remedied in Section 6.2, where we will exhibit particular subclasses that do enjoy the SPASs property.

4.1 Diagonal functors

In the introduction we gave a simple example of a pathological locally covariant theory with target category Alg. As we now show, this example may be placed within a more general setting, which provides a broader class of pathological theories and enables the consideration of more general categories for Phys.

We begin with a simple categorical construction. Given any two categories C and C', the functors between C and C' form the objects of a category $\mathrm{Funct}(\mathsf{C},\mathsf{C}')$ (also written C'^{C} in the literature) in which morphisms are natural transformations between pairs of functors. In particular, this applies to the locally covariant theories, which (in the BFV definition) are precisely the objects of $\mathsf{LCT}_0 = \mathrm{Funct}(\mathsf{Loc}_0,\mathsf{Phys})$. Iterating this construction, we may also consider functors from C to $\mathrm{Funct}(\mathsf{C},\mathsf{C}')$; any such functor then induces a functor in $\mathrm{Funct}(\mathsf{C},\mathsf{C}')$ by the following 'diagonal construction'.

Proposition 4.2 *Given $\varphi \in \mathrm{Funct}(\mathsf{C},\mathrm{Funct}(\mathsf{C},\mathsf{C}'))$, define maps of objects $A \in \mathsf{C}$ and morphisms $f \in \mathsf{C}(A,B)$ of C to objects and morphisms of C' by*

$$\varphi_\Delta(A) = \varphi(A)(A)$$
$$\varphi_\Delta(f) = \varphi(f)_B \circ \varphi(A)(f).$$

[14] The question of whether there is a sensible notion of 'the same theory' in spacetimes of *different* dimensions is an interesting one, to which we hope to return elsewhere.

Then φ_Δ is a functor from C to C'; we refer to φ_Δ as the diagonal *of φ. Moreover, if φ, φ' are elements of* $\mathrm{Funct}(\mathsf{C}, \mathrm{Funct}(\mathsf{C}, \mathsf{C}'))$ *and $\xi : \varphi \dot{\to} \varphi'$ is natural, there is a natural transformation $\xi_\Delta : \varphi_\Delta \dot{\to} \varphi'_\Delta$ with components $(\xi_\Delta)_A = (\xi_A)_A$. The map $\xi \mapsto \xi_\Delta$ is in fact a functor from* $\mathrm{Funct}(\mathsf{C}, \mathrm{Funct}(\mathsf{C}, \mathsf{C}'))$ *to* $\mathrm{Funct}(\mathsf{C}, \mathsf{C}')$.

Remarks: (1) The expressions above are well-defined because $\varphi(f) : \varphi(A) \dot{\to} \varphi(B)$; diagrammatically, $\varphi_\Delta(f)$ is the diagonal of the naturality square

$$\begin{tikzcd}
A \arrow[d, "f"'] & \varphi(A)(A) \arrow[r, "\varphi(f)_A"] \arrow[d, "\varphi(A)(f)"'] \arrow[dr, "\varphi_\Delta(f)"] & \varphi(B)(A) \arrow[d, "\varphi(B)(f)"] \\
B & \varphi(A)(B) \arrow[r, "\varphi(f)_B"'] & \varphi(B)(B)
\end{tikzcd}$$

and we also have $\varphi_\Delta(f) = \varphi(B)(f) \circ \varphi(f)_A$.
(2) Given any functor $F : \mathsf{C} \to \mathsf{C}'$, let φ be the constant functor $\mathsf{C} \to \mathrm{Funct}(\mathsf{C}, \mathsf{C}')$ taking the value F on all objects. Then $F = \varphi_\Delta$.
Proof: As φ and $\varphi(A)$ are functors, we have

$$\varphi_\Delta(\mathrm{id}_A) = \varphi(\mathrm{id}_A)_A \circ \varphi(A)(\mathrm{id}_A) = (\mathrm{id}_{\varphi(A)})_A \circ \mathrm{id}_{\varphi(A)(A)} = \mathrm{id}_{\varphi(A)(A)} = \mathrm{id}_{\varphi_\Delta(A)}$$

for any $A \in \mathsf{C}$. Moreover, if $g : B \to C$, we have

$$\varphi_\Delta(g \circ f) = \varphi(C)(g \circ f) \circ \varphi(g \circ f)_A = \varphi(C)(g) \circ \underbrace{\varphi(C)(f) \circ \varphi(g)_A}_{= \varphi(g)_B \circ \varphi(B)(f)} \circ \varphi(f)_A = \varphi_\Delta(g) \circ \varphi_\Delta(f),$$

in which we have used the naturality of $\varphi(g) : \varphi(B) \dot{\to} \varphi(C)$.

Now suppose that $\xi : \varphi \dot{\to} \varphi'$. Noting that $\xi_A : \varphi(A) \to \varphi'(A)$ is itself a natural transformation, each $(\xi_A)_A$ is a morphism from $(\xi_A)_A : \varphi(A)(A) \to \varphi'(A)(A)$. Given $f : A \to B$ we compute

$$\begin{aligned}
\varphi'_\Delta(f) \circ (\xi_\Delta)_A &= \varphi'(B)(f) \circ \varphi'(f)_A \circ (\xi_A)_A = \varphi'(B)(f) \circ (\varphi'(f) \circ \xi_A)_A \\
&= \varphi'(B)(f) \circ (\xi_B \circ \varphi(f))_A = \varphi'(B)(f) \circ (\xi_B)_A \circ \varphi(f)_A \\
&= (\xi_B)_B \circ \varphi(B)(f) \circ \varphi(f)_A = (\xi_\Delta)_B \circ \varphi_\Delta(f),
\end{aligned}$$

thus establishing naturality. (The above computation may be displayed diagrammatically using a commuting cube). It is simple to check the functor property and we skip the proof. □

In our examples, it will be convenient to construct functors from the category of spacetimes to the category of locally covariant theories using a construction of the following type.

Lemma 4.3 *Let* C *and* C' *be categories and* I *be a partially ordered set, which we may regard as a category [with a single arrow* $\iota \to \iota'$ *if and only if* $\iota \preccurlyeq \iota'$*], and suppose a functor* $\beta : \mathsf{I} \to \mathrm{Funct}(\mathsf{C}, \mathsf{C}')$ *is given. Then every functor* $\lambda : \mathsf{C} \to \mathsf{I}$ *determines a functor* $\varphi = \beta \circ \lambda : \mathsf{C} \to \mathrm{Funct}(\mathsf{C}, \mathsf{C}')$ *and hence a diagonal functor* $\varphi_\Delta \in \mathrm{Funct}(\mathsf{C}, \mathsf{C}')$*. Moreover, any natural transformation* $\zeta : \lambda \stackrel{.}{\to} \lambda'$ *between* $\lambda, \lambda' \in \mathrm{Funct}(\mathsf{C}, \mathsf{I})$ *induces a natural transformation between the corresponding diagonal functors.*

Remark: A functor $\lambda : \mathsf{C} \to \mathsf{I}$ is equivalent to labelling each object A of C with an element $\lambda(A) \in \mathsf{I}$, subject to the requirement that $\lambda(A) \preccurlyeq \lambda(B)$ if there is a C-morphism from A to B. The existence of a natural transformation between λ and λ' amounts to the condition that $\lambda(A) \preccurlyeq \lambda'(A)$ for all A. The use of partially ordered sets is simply for convenience and familiarity.
Proof: Given functors β and λ as described, it is obvious that $\varphi = \beta \circ \lambda \in \mathrm{Funct}(\mathsf{C}, \mathrm{Funct}(\mathsf{C}, \mathsf{C}'))$. Given $\zeta : \lambda \stackrel{.}{\to} \lambda'$, the maps $\xi_A = \beta(\zeta_A)$ form the components of a natural $\xi : \varphi \stackrel{.}{\to} \varphi'$ by functoriality of β. Hence $\xi_\Delta : \varphi_\Delta \stackrel{.}{\to} \varphi'_\Delta$ has components $(\xi_\Delta)_A = \beta(\zeta_A)_A$. $\square$

4.2 Diagonal theories in LCT_0 and LCT

Any functor $\varphi : \mathsf{Loc} \to \mathsf{LCT}$ assigns to each spacetime $\boldsymbol{M} \in \mathsf{Loc}$ a locally covariant theory defined on *all* spacetimes, i.e., a functor $\varphi(\boldsymbol{M}) : \mathsf{Loc} \to \mathsf{Phys}$, and assigns to each embedding $\psi : \boldsymbol{M} \to \boldsymbol{N}$ a natural transformation $\varphi(\psi) : \varphi(\boldsymbol{M}) \stackrel{.}{\to} \varphi(\boldsymbol{N})$ between the theories assigned to $\boldsymbol{M}$ and $\boldsymbol{N}$ respectively. The diagonal functor φ_Δ is then again an object of LCT and hence a theory in its own right; we will refer to it as a *diagonal theory*.

By Remark (2) following Prop. 4.2, every theory $\mathscr{A} \in \mathsf{LCT}_0$ is a diagonal theory in which $\varphi : \mathsf{Loc}_0 \to \mathsf{Phys}$ is a constant functor taking the value $\mathscr{A}$ in all spacetimes. So diagonal theories certainly exist. Our aim in this subsection is to investigate some of the general properties of diagonal theories and to develop criteria that would give various types of desirable or pathological properties; in particular, that violations of SPASs can, in principle, be achieved with theories that are otherwise well-behaved. In the following subsection we will show that such diagonal theories exist under fairly mild restrictions on the category Phys.

Our discussion is expressed for diagonal theories in LCT; all our remarks in this subsection apply equally to diagonal theories in LCT_0 on replacing Loc by Loc_0, LCT by LCT_0, and $\mathscr{O}(\boldsymbol{M})$ by $\mathscr{O}_0(\boldsymbol{M})$.

The kinematic net If $\boldsymbol{M} \in \mathsf{Loc}$ and $O \in \mathscr{O}(\boldsymbol{M})$, the kinematic local algebra is

$$\begin{aligned}(\varphi_\Delta)^{\mathrm{kin}}_{\boldsymbol{M};O} = \varphi_\Delta(\iota_{\boldsymbol{M};O}) = \varphi(\boldsymbol{M})(\iota_{\boldsymbol{M};O}) \circ \varphi(\iota_{\boldsymbol{M};O})_{\boldsymbol{M}|_O} = \varphi(\boldsymbol{M})^{\mathrm{kin}}_{\boldsymbol{M};O} \circ \varphi(\iota_{\boldsymbol{M};O})_{\boldsymbol{M}|_O} \\ \le \varphi(\boldsymbol{M})^{\mathrm{kin}}_{\boldsymbol{M};O}.\end{aligned}$$

If there exists any morphism $\psi : \boldsymbol{L} \to \boldsymbol{M}$ such that $\varphi(\psi)_{\boldsymbol{M}}$ is not an isomorphism then $(\varphi_\Delta)^{\mathrm{kin}}_{\boldsymbol{M};\psi(\boldsymbol{L})}$ is a proper subobject of $\varphi(\boldsymbol{M})^{\mathrm{kin}}_{\boldsymbol{M};\psi(\boldsymbol{L})}$.

The timeslice property Suppose $\varphi : \mathsf{Loc} \to \mathsf{LCT}$. For any morphism $\psi : \boldsymbol{M} \to \boldsymbol{N}$ we have $\varphi_\Delta(\psi) = \varphi(\psi)_{\boldsymbol{N}} \circ \varphi(\boldsymbol{M})(\psi)$. Accordingly, a *sufficient* condition for φ_Δ to satisfy the timeslice property is that both the following hold: (i) for every $\boldsymbol{M} \in \mathsf{Loc}$, $\varphi(\boldsymbol{M})$ satisfies the timeslice property and (ii) φ obeys the timeslice property in that $\varphi(\psi)$ is a natural isomorphism whenever ψ is Cauchy.

In particular, suppose that $\varphi = \beta \circ \lambda$, where I is a poset (regarded as a category) and β and λ are functors. Then the sufficient condition just mentioned becomes (i) for each $\ell \in \operatorname{Im} \lambda$, $\beta(\ell)$ obeys the timeslice axiom, and (ii) λ is constant on Cauchy-wedge-connected components of Loc. To see this, note that (ii) implies that if $\psi : \boldsymbol{M} \to \boldsymbol{N}$ is Cauchy, then $\lambda(\psi) = \mathrm{id}_{\lambda(\boldsymbol{M})}$ and hence $\varphi(\psi) = \mathrm{id}_{\varphi(\boldsymbol{M})}$.

The relative Cauchy evolution Suppose $\varphi : \mathsf{Loc} \to \mathsf{LCT}$ is such that every $\varphi(\boldsymbol{M})$ obeys the timeslice axiom and so does φ_Δ.

Lemma 4.4 *If $\boldsymbol{M} \stackrel{\psi}{\to} \boldsymbol{N}$ is a Cauchy morphism then $\varphi(\psi)_{\boldsymbol{M}}$ and $\varphi(\psi)_{\boldsymbol{N}}$ are isomorphisms.*

Proof: As $\varphi(\boldsymbol{M})$ and $\varphi(\boldsymbol{N})$ obey the timeslice axiom, $\varphi(\boldsymbol{M})(\psi)$ and $\varphi(\boldsymbol{N})(\psi)$ are isomorphisms. As $\varphi_\Delta(\psi)$ is also an isomorphism, the result follows because $\varphi(\psi)_{\boldsymbol{M}} = (\varphi(\boldsymbol{N})(\psi))^{-1} \circ \varphi_\Delta(\psi)$ and $\varphi(\psi)_{\boldsymbol{N}} = \varphi_\Delta(\psi) \circ (\varphi(\boldsymbol{M})(\psi))^{-1}$. $\square$

Proposition 4.5 *For any $\boldsymbol{h} \in H(\boldsymbol{M})$ we have*

$$\tau_{\boldsymbol{M}}^{(\varphi_\Delta)\pm}[\boldsymbol{h}] = \tau_{\boldsymbol{M}}^{(\varphi(\boldsymbol{M}[\boldsymbol{h}]))\pm}[\boldsymbol{h}] \circ \varphi(\jmath_{\boldsymbol{M}}^{\pm}[\boldsymbol{h}])_{\boldsymbol{M}} \circ \left(\varphi(\imath_{\boldsymbol{M}}^{\pm}[\boldsymbol{h}])_{\boldsymbol{M}}\right)^{-1}, \tag{4.1}$$

where $\boldsymbol{M} \stackrel{\imath_{\boldsymbol{M}}^{\pm}[\boldsymbol{h}]}{\longleftarrow} \boldsymbol{M}^{\pm}[\boldsymbol{h}] \stackrel{\jmath_{\boldsymbol{M}}^{\pm}[\boldsymbol{h}]}{\longrightarrow} \boldsymbol{M}[\boldsymbol{h}]$ are the future and past Cauchy wedges induced by $\boldsymbol{h}$. Hence

$$\mathrm{rce}_{\boldsymbol{M}}^{(\varphi_\Delta)}[\boldsymbol{h}] = \varphi(\imath_{\boldsymbol{M}}^{-}[\boldsymbol{h}])_{\boldsymbol{M}} \circ \varphi(\jmath_{\boldsymbol{M}}^{-}[\boldsymbol{h}])_{\boldsymbol{M}}^{-1} \circ \varphi(\jmath_{\boldsymbol{M}}^{+}[\boldsymbol{h}])_{\boldsymbol{M}} \circ \left(\varphi(\imath_{\boldsymbol{M}}^{+}[\boldsymbol{h}])_{\boldsymbol{M}}\right)^{-1} \circ \mathrm{rce}_{\boldsymbol{M}}^{(\varphi(\boldsymbol{M}))}[\boldsymbol{h}].$$

If φ also obeys timeslice [i.e., $\varphi(\psi)$ is a natural isomorphism for each Cauchy morphism ψ] then these results may be written more compactly as

$$\tau_{\boldsymbol{M}}^{(\varphi_\Delta)\pm}[\boldsymbol{h}] = \tau_{\boldsymbol{M}}^{(\varphi(\boldsymbol{M}[\boldsymbol{h}]))\pm}[\boldsymbol{h}] \circ (\tau_{\boldsymbol{M}}^{(\varphi)\pm}[\boldsymbol{h}])_{\boldsymbol{M}}$$

and

$$\mathrm{rce}_{\boldsymbol{M}}^{(\varphi_\Delta)}[\boldsymbol{h}] = (\mathrm{rce}_{\boldsymbol{M}}^{\varphi}[\boldsymbol{h}])_{\boldsymbol{M}} \circ \mathrm{rce}_{\boldsymbol{M}}^{(\varphi(\boldsymbol{M}))}[\boldsymbol{h}].$$

Proof: As usual, $\tau_{\boldsymbol{M}}^{(\varphi_\Delta)\pm}[\boldsymbol{h}]$ is the unique morphism such that $\tau_{\boldsymbol{M}}^{(\varphi_\Delta)\pm}[\boldsymbol{h}] \circ \varphi_\Delta(\imath_{\boldsymbol{M}}^{\pm}[\boldsymbol{h}]) = \varphi_\Delta(\jmath_{\boldsymbol{M}}^{\pm}[\boldsymbol{h}])$, i.e.,

$$\begin{aligned}
\tau_{\boldsymbol{M}}^{(\varphi_\Delta)\pm}[\boldsymbol{h}] \circ \varphi(\imath_{\boldsymbol{M}}^{\pm}[\boldsymbol{h}])_{\boldsymbol{M}} \circ \varphi(\boldsymbol{M}^{\pm})(\imath_{\boldsymbol{M}}^{\pm}[\boldsymbol{h}]) &= \varphi(\jmath_{\boldsymbol{M}}^{\pm}[\boldsymbol{h}])_{\boldsymbol{M}[\boldsymbol{h}]} \circ \varphi(\boldsymbol{M}^{\pm})(\jmath_{\boldsymbol{M}}^{\pm}[\boldsymbol{h}]) \\
&= \varphi(\jmath_{\boldsymbol{M}}^{\pm}[\boldsymbol{h}])_{\boldsymbol{M}[\boldsymbol{h}]} \circ \tau_{\boldsymbol{M}}^{(\varphi(\boldsymbol{M}^{\pm}))\pm}[\boldsymbol{h}] \circ \varphi(\boldsymbol{M}^{\pm})(\imath_{\boldsymbol{M}}^{\pm}[\boldsymbol{h}]).
\end{aligned}$$

As $\varphi(\boldsymbol{M}^{\pm})(\imath^{\pm}_{\boldsymbol{M}}[\boldsymbol{h}])$ and (by Lemma 4.4) $\varphi(\imath^{\pm}_{\boldsymbol{M}}[\boldsymbol{h}])_{\boldsymbol{M}}$ are isomorphisms, Eq. (4.1) holds. Accordingly,

$$
\begin{aligned}
\mathrm{rce}^{(\varphi_\Delta)}_{\boldsymbol{M}}[\boldsymbol{h}] &= \left(\tau^{(\varphi_\Delta)-}_{\boldsymbol{M}}[\boldsymbol{h}]\right)^{-1} \circ \tau^{(\varphi_\Delta)+}_{\boldsymbol{M}}[\boldsymbol{h}] \\
&= \varphi(\imath^{-}_{\boldsymbol{M}}[\boldsymbol{h}])_{\boldsymbol{M}} \circ \left(\varphi(\jmath^{-}_{\boldsymbol{M}}[\boldsymbol{h}])_{\boldsymbol{M}}\right)^{-1} \circ \mathrm{rce}^{(\varphi(\boldsymbol{M}[\boldsymbol{h}]))}_{\boldsymbol{M}}[\boldsymbol{h}] \circ \varphi(\jmath^{+}_{\boldsymbol{M}}[\boldsymbol{h}])_{\boldsymbol{M}} \circ \left(\varphi(\imath^{+}_{\boldsymbol{M}}[\boldsymbol{h}])_{\boldsymbol{M}}\right)^{-1} \\
&= \varphi(\imath^{-}_{\boldsymbol{M}}[\boldsymbol{h}])_{\boldsymbol{M}} \circ \left(\varphi(\jmath^{-}_{\boldsymbol{M}}[\boldsymbol{h}])_{\boldsymbol{M}}\right)^{-1} \circ \varphi(\jmath^{+}_{\boldsymbol{M}}[\boldsymbol{h}])_{\boldsymbol{M}} \circ \mathrm{rce}^{(\varphi(\boldsymbol{M}^{+}))}_{\boldsymbol{M}}[\boldsymbol{h}] \circ \left(\varphi(\imath^{+}_{\boldsymbol{M}}[\boldsymbol{h}])_{\boldsymbol{M}}\right)^{-1} \\
&= \varphi(\imath^{-}_{\boldsymbol{M}}[\boldsymbol{h}])_{\boldsymbol{M}} \circ \left(\varphi(\jmath^{-}_{\boldsymbol{M}}[\boldsymbol{h}])_{\boldsymbol{M}}\right)^{-1} \circ \varphi(\jmath^{+}_{\boldsymbol{M}}[\boldsymbol{h}])_{\boldsymbol{M}} \circ \left(\varphi(\imath^{+}_{\boldsymbol{M}}[\boldsymbol{h}])_{\boldsymbol{M}}\right)^{-1} \circ \mathrm{rce}^{(\varphi(\boldsymbol{M}))}_{\boldsymbol{M}}[\boldsymbol{h}]
\end{aligned}
$$

as required, where we have used Prop. 3.8 in the last two steps. The remaining statements are straightforward. □

It is clear from the above result that the diagonal theory φ_Δ does not necessarily have the same relative Cauchy evolution in spacetime $\boldsymbol{M}$ as $\varphi(\boldsymbol{M})$. In principle, this allows the stress–energy tensor to have a component that reflects the dynamics of the functor φ as well as the dynamics of the theory in spacetime $\boldsymbol{M}$. We do not know whether this can be realised in actual examples, however. Certainly, if φ factors through a poset, then (as we have already seen) φ maps any Cauchy morphism to an identity and so we have the simpler formulae

$$
\tau^{(\varphi_\Delta)\pm}_{\boldsymbol{M}}[\boldsymbol{h}] = \tau^{(\varphi(\boldsymbol{M}[\boldsymbol{h}]))\pm}_{\boldsymbol{M}}[\boldsymbol{h}] \tag{4.2}
$$

$$
\mathrm{rce}^{(\varphi_\Delta)}_{\boldsymbol{M}}[\boldsymbol{h}] = \mathrm{rce}^{(\varphi(\boldsymbol{M}))}_{\boldsymbol{M}}[\boldsymbol{h}]. \tag{4.3}
$$

Any diagonal theory in which these relations hold will be described as *ordinary*.

Comparison of theories and failure of SPASs in LCT Suppose $\varphi = \beta \circ \lambda$, where $\lambda : \mathsf{Loc} \to \mathsf{I}$ and $\beta : \mathsf{I} \to \mathsf{LCT}$ with I a poset. Suppose that there are $\ell, \ell' \in \mathsf{I}$ such that $\ell \preccurlyeq \lambda(\boldsymbol{M}) \preccurlyeq \ell'$ for all $\boldsymbol{M} \in \mathsf{Loc}$, with both ℓ and ℓ' being attained on certain spacetimes, and assume that $\beta(\ell, \ell')$ is not an isomorphism. By the remark following Lem. 4.3, this gives natural transformations $\kappa_\ell \stackrel{.}{\to} \lambda \stackrel{.}{\to} \kappa_{\ell'}$, where κ_p is the constant functor taking the value p on all objects; hence, by Lem. 4.3, there are natural transformations

$$
\beta(\ell) = (\beta \circ \kappa_\ell)_\Delta \stackrel{.}{\to} (\beta \circ \lambda)_\Delta \stackrel{.}{\to} (\beta \circ \kappa_{\ell'})_\Delta = \beta(\ell') \tag{4.4}
$$

whose components in an arbitrary spacetime $\boldsymbol{M}$ are

$$
\beta(\ell)(\boldsymbol{M}) \xrightarrow{\beta(\ell, \lambda(\boldsymbol{M}))_{\boldsymbol{M}}} (\beta \circ \lambda)_\Delta(\boldsymbol{M}) \xrightarrow{\beta(\lambda(\boldsymbol{M}), \ell')_{\boldsymbol{M}}} \beta(\ell')(\boldsymbol{M}),
$$

composing to $\beta(\ell, \ell')_{\boldsymbol{M}}$. Accordingly the two naturals in Eq. (4.4) compose to $\beta(\ell, \ell')$.

Now let $\boldsymbol{L}$ and $\boldsymbol{L}'$ be spacetimes with $\lambda(\boldsymbol{L}) = \ell$, $\lambda(\boldsymbol{L}') = \ell'$. Then the first natural is an identity in spacetime $\boldsymbol{L}$, while the second is an identity in spacetime $\boldsymbol{L}'$. Thus both are partial isomorphisms. If the SPASs property were to hold on (any class of theories including) $\beta(\ell)$, $(\beta \circ \lambda)_\Delta$ and $\beta(\ell')$, then both naturals would have to be isomorphisms, which contradicts the fact that their composite, $\beta(\ell, \ell')$, is not an isomorphism.

In particular, if one or both of the theories $\beta(\ell)$ and $\beta(\ell')$ are regarded as individually representing the same physics in all spacetimes (by some reasonable definition) then it is clearly impossible for $(\beta \circ \lambda)_\Delta$ to represent the same physics in all spacetimes (by the same definition).

This discussion shows that the failure of SPASs can be exhibited quite straightforwardly, given suitable functors β and λ. In the next subsection, we will give some concrete constructions which achieve this goal. We have presented the discussion so far in fairly abstract terms, partly to facilitate discussion of general categories Phys and partly because a wide range of constructions can be given and we wish to emphasise that the issue runs more deeply than a few isolated counterexamples (each of which, perhaps, could be removed by some *ad hoc* additional assumptions). In addition, it may be that diagonal theories may provide useful examples in other contexts, e.g., locally covariant theories that do not obey the timeslice axiom.

4.3 Specific Examples

To start, let us consider the problem of constructing a functor from Loc_0 to a poset. There are many ways of doing this, and the reader should regard the examples presented here as indicative rather than exhaustive.

For a first example, fix a constant $R_0 > 0$ with dimensions of length^{-2} and define

$$\lambda(\boldsymbol{M}) = \begin{cases} 2 & \sup R_{\boldsymbol{M}} > R_0 \\ 1 & \sup R_{\boldsymbol{M}} \le R_0, \end{cases} \tag{4.5}$$

where $R_{\boldsymbol{M}}$ is the scalar curvature on $\boldsymbol{M} \in \mathsf{Loc}_0$ and the supremum is taken over all of $\boldsymbol{M}$. It is clear that if $\psi : \boldsymbol{M} \to \boldsymbol{N}$ then $\lambda(\boldsymbol{M}) \le \lambda(\boldsymbol{N})$, so λ is indeed a functor from Loc_0 to N, i.e., the natural numbers with their usual ordering.

This particular functor is not constant on Cauchy-wedge-connected components of Loc_0, however. To see this, consider a spacetime containing Cauchy surfaces Σ_1 and Σ_2 so that the scalar curvature exceeds R_0 near Σ_1, but is everywhere less than R_0 in a globally hyperbolic neighbourhood of Σ_2. This induces a Cauchy wedge connecting a spacetime with $\lambda = 1$ to a spacetime where $\lambda = 2$. Thus diagonal theories based on such functors would not be expected to have the timeslice property. However, we will find a use for this example below.

A different type of example is constructed by choosing any function $\mu : \mathsf{Loc}_0 \to \mathbb{N}$ such that (i) $\mu(\boldsymbol{M})$ depends only on the oriented-diffeomorphism class of the smooth spacelike Cauchy surfaces of $\boldsymbol{M}$; (ii) μ takes its minimum value on all spacetimes with noncompact Cauchy surfaces. This is obviously constant on Cauchy-wedge-connected components by Prop. 2.4. To see that it is a functor from Loc_0 to N, we take any morphism $\psi : \boldsymbol{M} \to \boldsymbol{N}$ in Loc_0. If $\boldsymbol{M}$ has noncompact Cauchy surface, then $\mu(\boldsymbol{M}) \le \mu(\boldsymbol{N})$ by condition (ii). If, on the other hand, $\boldsymbol{M}$ has compact Cauchy surfaces, then Prop. 2.3(a) entails that $\boldsymbol{M}$ and $\boldsymbol{N}$ have oriented-diffeomorphic Cauchy surfaces and hence $\mu(\boldsymbol{M}) = \mu(\boldsymbol{N})$. Thus

$\mu \in \mathrm{Funct}(\mathsf{Loc}_0, \mathsf{N})$. [Equally, this construction gives a functor to $\mathrm{Im}\,\mu$, equipped with the partial ordering in which $p \preccurlyeq q$ iff $p = q$ or $p = \min \mathrm{Im}\,\mu$].

In view of the comments in the previous subsection, diagonal theories $(\beta \circ \mu)_\Delta$ will obey the timeslice property provided that $\beta(\ell)$ obeys timeslice for each $\ell \in \mathrm{Im}\,\mu$.

Turning to the case of possibly disconnected spacetimes, one way of constructing a functor from Loc to N is to take any functor $\lambda_0 : \mathsf{Loc}_0 \to \mathsf{N}$ and to define

$$\lambda(\boldsymbol{M}) = \max_{\boldsymbol{C} \in \mathrm{Cpts}(\boldsymbol{M})} \lambda_0(\boldsymbol{C})$$

for $\boldsymbol{M} \in \mathsf{Loc}$. Consider any Loc-morphism $\psi : \boldsymbol{M} \to \boldsymbol{N}$ and let $\boldsymbol{B}$ be a component of $\boldsymbol{M}$ such that $\lambda(\boldsymbol{M}) = \lambda_0(\boldsymbol{B})$. Then there is a component $\boldsymbol{C}$ of $\boldsymbol{N}$ so that $\psi(\boldsymbol{B}) \subset \boldsymbol{C}$ and a Loc_0-morphism $\psi^{\boldsymbol{B}}_{\boldsymbol{C}} : \boldsymbol{B} \to \boldsymbol{C}$. Then

$$\lambda(\boldsymbol{M}) = \lambda_0(\boldsymbol{B}) \le \lambda_0(\boldsymbol{C}) \le \lambda(\boldsymbol{N}),$$

which suffices to show that $\lambda \in \mathrm{Funct}(\mathsf{Loc}, \mathsf{N})$. Moreover, λ will be constant on Cauchy-wedge-connected components of Loc if λ_0 is constant on Cauchy-wedge-connected components of Loc_0.

There are many other possibilities. Let Surf be the set of smooth connected *compact* orientable $(n-1)$-manifolds modulo oriented-diffeomorphisms (n being the spacetime dimension). To every $\boldsymbol{M} \in \mathsf{Loc}$ there is a function $\nu_{\boldsymbol{M}} : \mathrm{Surf} \to \mathbb{N}_0$ such that $\nu_{\boldsymbol{M}}(\Sigma)$ is the number of connected components of $\boldsymbol{M}$ whose Cauchy surfaces are oriented-diffeomorphic to Σ. Evidently $\nu_{\boldsymbol{M}}(\Sigma)$ is nonzero for at most finitely many $\Sigma \in \mathrm{Surf}$; using Prop. 2.3(b) it is easily seen that the existence of a morphism $\psi : \boldsymbol{M} \to \boldsymbol{N}$ entails that $\nu_{\boldsymbol{M}}(\Sigma) \le \nu_{\boldsymbol{N}}(\Sigma)$ for all $\Sigma \in \mathrm{Surf}$ (we are only counting compact connected components). A wide variety of functors $\lambda : \mathsf{Loc} \to \mathsf{N}$ may now be constructed, such as

$$\lambda(\boldsymbol{M}) = a + \sum_{\Sigma} m(\Sigma) \nu_{\boldsymbol{M}}(\Sigma)^{p(\Sigma)}$$

for $a \in \mathbb{N}$ and any functions $m, p : \mathrm{Surf} \to \mathbb{N}_0$. All such functors are constant on Cauchy-wedge-connected components of Loc, because Cauchy-wedge-connected spacetimes $\boldsymbol{M}$ and $\boldsymbol{N}$ have oriented-diffeomorphic Cauchy surfaces, so the functions $\nu_{\boldsymbol{M}}$ and $\nu_{\boldsymbol{N}}$ coincide.

We have shown that it is possible to construct functors from Loc_0 and Loc to various posets in various ways. There are also various ways of obtaining functors from a poset to LCT as shown by the following examples (all of which adapt straightforwardly to LCT_0):

1. If I is the poset $\mathbb{N}$ with the ordering $p \preccurlyeq q$ iff $p = 1$ or $p = q$, we may proceed by setting $\beta(1) = \mathscr{I}$, the initial theory, and choose $\beta(p) \in \mathsf{LCT}$ arbitrarily for $p \ge 2$. To the arrow $1 \to p$ assign the natural $\mathscr{I}_{\beta(p)} : \mathscr{I} \stackrel{.}{\to} \beta(p)$ that arises because $\mathscr{I}$ is initial. All other arrows in I are identities, and we assign to each id_p the morphism $\mathrm{id}_{\beta(p)}$ [evidently this is compatible with the previous assignment for $p = 1$]. Then $\beta \in \mathrm{Funct}(\mathsf{I}, \mathsf{LCT})$.

2. Suppose Phys admits an endofunctor $\mathscr{F}$ and a natural $\eta : \mathscr{F} \stackrel{.}{\to} \mathrm{id}_{\mathsf{Phys}}$.[15] Given any $\mathscr{A} \in \mathsf{LCT}$ there is a functor $\beta : (\{1,2\}, \leq) \to \mathsf{LCT}$ with

$$\beta(1) = \mathscr{F} \circ \mathscr{A}, \quad \beta(2) = \mathscr{A}, \qquad \beta(\mathrm{id}_1) = \mathrm{id}_{\mathscr{F}\circ\mathscr{A}}, \quad \beta(1 \to 2) = \eta, \quad \beta(\mathrm{id}_2) = \mathrm{id}_{\mathscr{A}}.$$

3. If Phys has a monoidal structure then, as discussed in Sect. 3.2, we obtain a functor $\beta : \mathsf{N} \to \mathsf{LCT}$ with $\beta(k) = \mathscr{A}^{\otimes k}$ and naturals $\beta(k,k') : \beta(k) \stackrel{.}{\to} \beta(k')$ for any $k \leq k'$.

Pursuing the third of these examples, let us suppose that $\mu_0 : \mathsf{Loc}_0 \to \mathbb{N}$ is constant on Cauchy-wedge-connected components of Loc_0, with $\mu_0(\boldsymbol{M}) = 1$ if $\boldsymbol{M}$ has noncompact Cauchy surfaces and $\mu_0(\boldsymbol{M}) \neq 1$ for some spacetimes. Let us suppose that the basic theory $\mathscr{A}$ has the timeslice property and is not idempotent, meaning that there is no $k \geq 2$ for which $\beta(1,k)$ is an isomorphism. Setting $\varphi = \beta \circ \lambda$, the φ_Δ is an ordinary diagonal theory in LCT_0, that will be denoted $\mathscr{A}^{[\mu_0]}$; it obeys the timeslice axiom because each $\mathscr{A}^{\otimes k}$ does.

In any spacetime $\boldsymbol{M}$, we have $\mathscr{A}^{[\mu_0]}(\boldsymbol{M}) = \mathscr{A}^{\otimes \mu_0(\boldsymbol{M})}(\boldsymbol{M})$; if $\psi : \boldsymbol{M} \to \boldsymbol{N}$ and $\mu_0(\boldsymbol{M}) \leq \mu_0(\boldsymbol{N})$ then

$$\mathscr{A}^{[\mu_0]}(\psi) = \beta(\mu_0(\boldsymbol{M}), \mu_0(\boldsymbol{N}))_{\boldsymbol{N}} \circ \mathscr{A}^{\otimes \mu_0(\boldsymbol{M})}(\psi) = \beta(\mu_0(\boldsymbol{M}), \mu_0(\boldsymbol{N}))_{\boldsymbol{N}} \circ \mathscr{A}(\psi)^{\otimes \mu_0(\boldsymbol{M})}$$

(if the category Phys is Alg, with the algebraic tensor product as the monoidal structure, then this has the action

$$\mathscr{A}^{[\mu_0]}(\psi)X = (\mathscr{A}^{\otimes \mu_0(\boldsymbol{M})}(\psi)X) \otimes \mathbf{1}_{\mathscr{A}(\boldsymbol{N})}^{\otimes(\mu_0(\boldsymbol{N}) - \mu_0(\boldsymbol{M}))}$$

on $X \in \mathscr{A}^{\otimes \mu_0(\boldsymbol{M})}(\boldsymbol{M})$). The kinematic net for φ_Δ produces subobjects $(\varphi_\Delta)^{\mathrm{kin}}_{\boldsymbol{M};O}$ that are proper subobjects of $\varphi(\boldsymbol{M})^{\mathrm{kin}}_{\boldsymbol{M};O}$ whenever $\mu_0(\boldsymbol{M}) > 1$ and $O \in \mathscr{O}_0(\boldsymbol{M})$ has noncompact Cauchy surface.

If, additionally, μ_0 is bounded with maximum value ℓ', then we may argue as in the previous subsection that the SPASs property cannot hold on any class of theories including $\mathscr{A}$, $\mathscr{A}^{[\mu_0]}$ and $\mathscr{A}^{\otimes \ell'}$; if either $\mathscr{A}$ or $\mathscr{A}^{\otimes \ell'}$ is regarded as representing the same physics in all spacetimes (by some definition), it follows that $\mathscr{A}^{[\mu_0]}$ cannot have this property (by the same definition).

This example is enough to show that LCT_0 will generally fail to have the SPASs property, except in the case that all its theories are idempotent. Similarly, in LCT, if we define $\mu(\boldsymbol{M}) = \max\{\mu_0(\boldsymbol{C}) : \boldsymbol{C} \in \mathrm{Cpts}(\boldsymbol{M})\}$, then the theory $\mathscr{A}^{[\mu]} := (\beta \circ \mu)_\Delta$ [with β now giving monoidal powers in LCT] has analogous properties and demonstrates the failure of SPASs in LCT.

We conclude this section by sketching two other examples to illustrate the range of bad behaviour that can occur. For the first, we return to the functor $\lambda : \mathsf{Loc}_0 \to \mathsf{N}$ of Eq. (4.5) and compose with the functor $\beta(k) = \mathscr{A}^{\otimes k}$, where $\mathscr{A}$ is nontrivial and has the timeslice property and is additive, in the sense that $\mathscr{A}(\boldsymbol{M})$ is generated by the $\mathscr{A}^{\mathrm{kin}}(\boldsymbol{M}; O_i)$ whenever the O_i form a cover of $\boldsymbol{M}$ by open globally hyperbolic spacetimes. The upshot

[15] See Sect. 3.2 for an example in TAlg.

is a theory $\mathscr{B} = (\beta \circ \lambda)_\Delta$ that coincides with $\mathscr{A}^{\otimes 2}$ in spacetimes whose scalar curvature somewhere exceeds R_0, and otherwise coincides with $\mathscr{A}$. (The theory $\mathscr{B}$ does not have the timeslice property.) Now consider a spacetime $\boldsymbol{M}$ that has a Cauchy surface Σ on which the scalar curvature is everywhere greater than R_0, but which also has an open globally hyperbolic region U on which the scalar curvature is everywhere less than R_0. Consider any cover $\boldsymbol{M} = \bigcup_i O_i$ by nonempty open globally hyperbolic spacetimes O_i. Then the O_i also cover Σ, and every O_i that intersects Σ nontrivially must have $\lambda(\boldsymbol{M}|_{O_i}) = 2$, so $\mathscr{B}^{\rm kin}(\boldsymbol{M}; O_i) = \mathscr{A}^{\otimes 2 {\rm kin}}(\boldsymbol{M}; O_i)$ for these particular regions. As $\mathscr{A}^{\otimes 2}$ has the timeslice property, this proves that $\mathscr{B}(\boldsymbol{M})$ is generated by the $\mathscr{B}^{\rm kin}(\boldsymbol{M}; O_i)$ with $O_i \cap \Sigma \neq 0$, and hence *a fortiori* by the full collection of $\mathscr{B}^{\rm kin}(\boldsymbol{M}; O_i)$. Thus the theory $\mathscr{B}$ is additive on $\boldsymbol{M}$ and has $\mathscr{B}(\boldsymbol{M}) = \mathscr{A}^{\otimes 2}(\boldsymbol{M})$; but at the same time, $\boldsymbol{M}$ contains a region U for which the local kinematic subobject $\beta^{\rm kin}_{\boldsymbol{M};U} = \beta(1,2) \circ \alpha^{\rm kin}_{\boldsymbol{M};U}$ corresponds to only *one* copy of the theory $\mathscr{A}$. This example stands as a counterpoint to the previous examples, where additivity would not be expected to hold in spacetimes with $\lambda = 2$.

Finally, as an extreme example, suppose Phys admits infinite monoidal products indexed over the naturals (i.e., a colimit of the functor giving finite monoidal powers). Then we may also form infinite powers $\mathscr{B}^{\otimes\infty}$ of any theory $\mathscr{B}$ in LCT. There is a right-shift endomorphism $\sigma : \mathscr{B}^{\otimes\infty} \stackrel{.}{\to} \mathscr{B}^{\otimes\infty}$ which is given (for $\mathsf{Phys} = \mathsf{Alg}$, say) by

$$\sigma_{\boldsymbol{M}} X = \mathbf{1}_{\mathscr{B}(\boldsymbol{M})} \otimes X,$$

which realises any such $\mathscr{B}^{\otimes\infty}$ as a proper subtheory of itself [i.e., σ is a non-automorphic endomorphism] except if $\mathscr{B}$ is the trivial theory. Now suppose $\mathscr{A}$ is any nontrivial locally covariant theory and let $\mathscr{B}$ be a diagonal theory that coincides with $\mathscr{I}$ in some spacetimes and $\mathscr{A}$ in others. Then the right-shift σ on $\mathscr{B}^{\otimes\infty}$ is a partial isomorphism, as $\sigma_{\boldsymbol{M}}$ is an isomorphism in every spacetime where $\mathscr{B}(\boldsymbol{M})$ is trivial. Of course, the theory $\mathscr{B}^{\otimes\infty}(\boldsymbol{M})$ is also trivial in such spacetimes, but by passing to the theory $\mathscr{A} \otimes \mathscr{B}^{\otimes\infty}$, we obtain a theory that is nontrivial in all spacetimes and admits an endomorphism ${\rm id}_{\mathscr{A}} \otimes \sigma$ that is a partial isomorphism but not an automorphism. Theories of this type cannot be regarded as obeying the same physics in all spacetimes by any reasonable notion: even the singleton $\{\mathscr{A} \otimes \mathscr{B}^{\otimes\infty}\}$ fails to have the SPASs property. One might suspect that theories admitting proper endomorphisms are always unphysical; elsewhere it will be shown that they conflict with natural requirements of nuclearity/energy compactness, which supports the idea that they must have infinitely many degrees of freedom available in bounded regions at finite energies [24].

We have described these examples in some detail to illustrate that a wide variety of bad behaviour can be exhibited by locally covariant theories. It seems likely that yet worse behaviour could be found.

5 Dynamical determination of local observables

5.1 The dynamical net

In Sect. 3.3, we saw how BFV used the functorial structure of a locally covariant theory to reconstruct a net structure of local observables. The idea was to regard the theory in a subregion of a spacetime as the theory assigned to that subregion when considered as a spacetime in its own right. We regard this as a kinematic description of the local physics. In this section we use the dynamics of the relative Cauchy evolution to give another description of local physics; the theory will be said to be *dynamically local* when these two descriptions of the local physics coincide. The diagonal theories, as we will see, include examples of theories that are not dynamically local; in [28] we will show that the Klein–Gordon theory is dynamically local both as a classical and a quantum theory (at nonzero mass; the massless case involves further subtleties).

To illustrate the general idea, suppose that Alg has been taken as the category Phys, and that $\mathscr{A}$ is a locally covariant theory in this setting. Fix a spacetime $\boldsymbol{M}$ and a compact set K therein. Any hyperbolic perturbation $\boldsymbol{h} \in H(\boldsymbol{M}; K^\perp)$ represents a modification in the spacetime in regions causally inaccessible from K; one would expect that observables localised within K should be insensitive to such changes. Taking this as a *definition* of what it means to be localised in K, we are led to study the subalgebra

$$\mathscr{A}^\bullet(\boldsymbol{M}; K) = \{A \in \mathscr{A}(\boldsymbol{M}) : \mathrm{rce}_{\boldsymbol{M}}[\boldsymbol{h}]A = A \text{ for all } \boldsymbol{h} \in H(\boldsymbol{M}; K^\perp)\}$$

as the candidate for the description of the local physics. Given an open globally hyperbolic subset with finitely many components (though not necessarily nonempty) $O \in \mathscr{O}(\boldsymbol{M})$ we may define the subalgebra $\mathscr{A}^{\mathrm{dyn}}(\boldsymbol{M}; O)$ of $\mathscr{A}(\boldsymbol{M})$ generated by the $\mathscr{A}^\bullet(\boldsymbol{M}; K)$ for a suitable class of compact subsets of O. (The simpler possibility of defining the $\mathscr{A}^\bullet(\boldsymbol{M}; \mathrm{cl}(O))$ as the local algebra of a relatively compact open globally hyperbolic set O would not generally give a match with the kinematic algebra $\mathscr{A}^{\mathrm{kin}}(\boldsymbol{M}; O)$ as can be seen in the example of the Klein–Gordon field [28]). To this end, for each nonempty $O \in \mathscr{O}(\boldsymbol{M})$ we define $\mathscr{K}(\boldsymbol{M}; O)$ to be the set of compact subsets contained in O and having a multi-diamond neighbourhood whose base is contained in O. In particular, this condition is obeyed by the empty set, so $\emptyset \in \mathscr{K}(\boldsymbol{M}; O)$ for all nonempty $O \in \mathscr{O}(\boldsymbol{M})$. By convention we also set $\mathscr{K}(\boldsymbol{M}; \emptyset) = \{\emptyset\}$. We use $\mathscr{K}(\boldsymbol{M})$ as a shorthand for $\mathscr{K}(\boldsymbol{M}; \mathcal{M})$.

This class is chosen for various reasons. The requirement to have a (multi)-diamond neighbourhood ensures, for example, that if $K \in \mathscr{K}(\boldsymbol{M})$ then $K^{\perp\perp}$ is again compact (see Lemma A.10; the proof relies on the relative compactness of multi-diamonds). We use multi-diamonds, rather than diamonds, to facilitate the treatment of sets O with more than one connected component; in some (but not all) theories one could insist on diamond neighbourhoods without loss. These issues will be discussed elsewhere.

We then define the *dynamical net* as the assignment to each $O \in \mathscr{O}(\boldsymbol{M})$ of the subalgebra

$$\mathscr{A}^{\mathrm{dyn}}(\boldsymbol{M}; O) = \bigvee_{K \in \mathscr{K}(\boldsymbol{M}; O)} \mathscr{A}^\bullet(\boldsymbol{M}; K) \tag{5.1}$$

in which the right-hand side denotes the Alg-subobject of $\mathscr{A}(\boldsymbol{M})$ generated by the $\mathscr{A}^\bullet(\boldsymbol{M};K)$ for $\mathscr{K}(\boldsymbol{M}) \ni K \subset O$. As $\emptyset \in \mathscr{K}(\boldsymbol{M};O)$, we always have $\mathscr{A}^\bullet(\boldsymbol{M};\emptyset) \subset \mathscr{A}^{\mathrm{dyn}}(\boldsymbol{M};O)$ for every O; in particular, $\mathscr{A}^{\mathrm{dyn}}(\boldsymbol{M};\emptyset) = \mathscr{A}^\bullet(\boldsymbol{M};\emptyset)$. As we will show in [28], Eq. (5.1) gives the correct local algebras for the simple model of the massive Klein–Gordon field.

More generally, the above ideas can be implemented in any category Phys satisfying our standing assumptions. As in the case of the kinematic net it is convenient to focus on the subobject morphisms; we will also find it useful to give 'universal' definitions for the various subobjects of interest.

Lemma 5.1 *For any compact subset K of $\boldsymbol{M}$ there exists a unique (up to isomorphism) subobject $\alpha^\bullet_{\boldsymbol{M};K}$ of $\mathscr{A}(\boldsymbol{M})$ such that (i)*

$$\mathrm{rce}_{\boldsymbol{M}}[\boldsymbol{h}] \circ \alpha^\bullet_{\boldsymbol{M};K} = \alpha^\bullet_{\boldsymbol{M};K} \qquad \forall \boldsymbol{h} \in H(\boldsymbol{M};K^\perp); \tag{5.2}$$

and (ii) if any other morphism α satisfies Eq. (5.2) in place of $\alpha^\bullet_{\boldsymbol{M};K}$, then $\alpha \le \alpha^\bullet_{\boldsymbol{M};K}$ in the subobject lattice of $\mathscr{A}(\boldsymbol{M})$.[16]

Proof: For each $\boldsymbol{h} \in H(\boldsymbol{M};K^\perp)$, let $\alpha_{\boldsymbol{h}}$ be the equaliser of $\mathrm{rce}_{\boldsymbol{M}}[\boldsymbol{h}]$ and $\mathrm{id}_{\mathscr{A}(\boldsymbol{M})}$ [which exists by assumption on Phys], i.e., a morphism such that $\mathrm{rce}_{\boldsymbol{M}}[\boldsymbol{h}] \circ \alpha_{\boldsymbol{h}} = \alpha_{\boldsymbol{h}}$, and so that any other morphism $\beta_{\boldsymbol{h}}$ obeying this equation in place of $\alpha_{\boldsymbol{h}}$ obeys $\beta_{\boldsymbol{h}} \le \alpha_{\boldsymbol{h}}$. Then any intersection

$$\alpha^\bullet_{\boldsymbol{M};K} \cong \bigwedge_{\boldsymbol{h}\in H(\boldsymbol{M};K^\perp)} \alpha_{\boldsymbol{h}}$$

(which exists by assumption on Phys) obeys Eq. (5.2): see, e.g., Lem. B.1. Any β also obeying this equation must in particular obey $\beta \le \alpha_{\boldsymbol{h}}$ for all $\boldsymbol{h} \in H(\boldsymbol{M};K^\perp)$ by the definition of the equaliser; accordingly, $\beta \le \alpha^\bullet_{\boldsymbol{M};K}$ by the definition of an intersection. $\square$

In the case $\mathsf{Phys} = \mathsf{Alg}$, $\alpha^\bullet_{\boldsymbol{M};K}$ is of course the inclusion morphism of $\mathscr{A}^\bullet(\boldsymbol{M};K)$ in $\mathscr{A}(\boldsymbol{M})$. Returning to the general case, Phys also has arbitrary categorical unions; accordingly, to each $O \in \mathscr{O}(\boldsymbol{M})$ there is a (unique up to isomorphism) subobject

$$\alpha^{\mathrm{dyn}}_{\boldsymbol{M};O} \cong \bigvee_{K\in\mathscr{K}(\boldsymbol{M};O)} \alpha^\bullet_{\boldsymbol{M};K} \tag{5.3}$$

(generalising the inclusion morphism of $\mathscr{A}^{\mathrm{dyn}}(\boldsymbol{M};O)$ in $\mathscr{A}(\boldsymbol{M})$ in the category Alg) that we take as the definition of the dynamical net. Denoting the domain of $\alpha^{\mathrm{dyn}}_{\boldsymbol{M};O}$ as $\mathscr{A}^{\mathrm{dyn}}(\boldsymbol{M};O)$, Eq. (5.3) means that (i) every $\alpha^\bullet_{\boldsymbol{M};K}$ (with $K \in \mathscr{K}(\boldsymbol{M};O)$) factorises (uniquely) via $\alpha^{\mathrm{dyn}}_{\boldsymbol{M};O}$ as $\alpha^\bullet_{\boldsymbol{M};K} = \alpha^{\mathrm{dyn}}_{\boldsymbol{M};O} \circ \alpha_{\boldsymbol{M};O;K}$; (ii) whenever there are morphisms β and γ and β_K such that $\beta \circ \beta_K = \gamma \circ \alpha^\bullet_{\boldsymbol{M};K}$ for every $K \in \mathscr{K}(\boldsymbol{M};O)$, there exists a unique $\xi : \mathscr{A}^{\mathrm{dyn}}(\boldsymbol{M};O) \to B$ such that

$$\beta_K = \xi \circ \alpha_{\boldsymbol{M};O;K} \qquad \text{and} \qquad \beta \circ \xi = \gamma \circ \alpha^{\mathrm{int}}_{\boldsymbol{M};O}$$

[16] Recall that this means there is a unique β such that $\alpha = \alpha^\bullet_{\boldsymbol{M};K} \circ \beta$.

for all $K \in \mathscr{K}(\boldsymbol{M}; O)$. Diagrammatically, fixing β and γ, if the outer portion of every diagram of the following form commutes as K varies in $\mathscr{K}(\boldsymbol{M}; O)$ then there is a unique ξ to make all the diagrams commute in full:

$$\begin{tikzcd}
 & \mathscr{A}^{\mathrm{dyn}}(\boldsymbol{M}; O) \arrow[r, "\alpha^{\mathrm{dyn}}_{\boldsymbol{M};O}"] \arrow[dd, dotted, "\xi"] & \mathscr{A}(\boldsymbol{M}) \arrow[dd, "\gamma"] \\
\mathscr{A}^{\bullet}(\boldsymbol{M}; K) \arrow[ur, "\alpha_{\boldsymbol{M};O;K}"] \arrow[dr, "\beta_K"'] & & \\
 & B \arrow[r, "\beta"] & C
\end{tikzcd}$$

(see Appendix B and [21] for more details on the union in general categories).

Although we have given notation for the domains of the morphisms $\alpha^{\bullet}_{\boldsymbol{M};K}$, $\alpha^{\mathrm{dyn}}_{\boldsymbol{M};O}$, one should bear in mind that it is the morphisms that are the significant entities. For the sake of familiarity we will write expressions such as $\mathscr{A}^{\bullet}(\boldsymbol{M}; K_1) \subset \mathscr{A}^{\bullet}(\boldsymbol{M}; K_2)$, but this must be understood as asserting that $\alpha^{\bullet}_{\boldsymbol{M};K_1}$ factorizes via $\alpha^{\bullet}_{\boldsymbol{M};K_2}$, i.e., $\alpha^{\bullet}_{\boldsymbol{M};K_1} = \alpha^{\bullet}_{\boldsymbol{M};K_2} \circ \beta$ for some $\beta : \mathscr{A}^{\bullet}(\boldsymbol{M}; K_1) \to \mathscr{A}^{\bullet}(\boldsymbol{M}; K_2)$. This is the order relation in the subobject lattice of $\mathscr{A}(\boldsymbol{M})$ (see, e.g., [21]). Similarly, $\mathscr{A}^{\bullet}(\boldsymbol{M}; K_1) \cong \mathscr{A}^{\bullet}(\boldsymbol{M}; K_2)$ asserts that $\alpha^{\bullet}_{\boldsymbol{M};K_1} = \alpha^{\bullet}_{\boldsymbol{M};K_2} \circ \beta$ with β an isomorphism, i.e., $\alpha^{\bullet}_{\boldsymbol{M};K_1} \cong \alpha^{\bullet}_{\boldsymbol{M};K_2}$ as subobjects. In the case of Alg or other category in which $\mathscr{A}^{\bullet}(\boldsymbol{M}; K)$ and $\mathscr{A}^{\mathrm{dyn}}(\boldsymbol{M}; O)$ are realised concretely as subsets of $\mathscr{A}(\boldsymbol{M})$, and the $\alpha^{\bullet}_{\boldsymbol{M};K}$, $\alpha^{\mathrm{dyn}}_{\boldsymbol{M};O}$ morphisms are set inclusions then the $\subset$ notation may be taken to indicate a subset and isomorphism can be upgraded to equality.

5.2 Properties of the dynamical net

The assignments $K \mapsto \mathscr{A}^{\bullet}(\boldsymbol{M}; K)$ and $O \mapsto \mathscr{A}^{\mathrm{dyn}}(\boldsymbol{M}; O)$ possess a number of properties that would be expected of a net of local algebras: namely, isotony, causal dynamics, and covariance with respect to isomorphisms.

Theorem 5.2 *(a) Suppose K_1, K_2 are compact and $J_{\boldsymbol{M}}(K_1) \subset J_{\boldsymbol{M}}(K_2)$ (in particular, if $K_1 \subset K_2$). Then $\mathscr{A}^{\bullet}(\boldsymbol{M}; K_1) \subset \mathscr{A}^{\bullet}(\boldsymbol{M}; K_2)$.*
(b) In consequence, we have

$$\mathscr{A}^{\bullet}(\boldsymbol{M}; K) \cong \mathscr{A}^{\bullet}(\boldsymbol{M}; K^{\perp\perp})$$

provided $K^{\perp\perp}$ is also compact (in particular, if $K \in \mathscr{K}(\boldsymbol{M})$) and, for any compact sets K_1, K_2,

$$\begin{aligned}
\mathscr{A}^{\bullet}(\boldsymbol{M}; K_1) \vee \mathscr{A}^{\bullet}(\boldsymbol{M}; K_2) &\subset \mathscr{A}^{\bullet}(\boldsymbol{M}; K_1 \cup K_2) \\
\mathscr{A}^{\bullet}(\boldsymbol{M}; K_1 \cap K_2) &\subset \mathscr{A}^{\bullet}(\boldsymbol{M}; K_1) \wedge \mathscr{A}^{\bullet}(\boldsymbol{M}; K_2)
\end{aligned}$$

and $\mathscr{A}^\bullet(\boldsymbol{M};\emptyset) \subset \mathscr{A}^\bullet(\boldsymbol{M};K)$ *for all compact* K*. (c) If* $\psi : \boldsymbol{M} \to \boldsymbol{N}$ *is an isomorphism then* $\mathscr{A}(\psi)$ *restricts to an isomorphism* $\mathscr{A}^\bullet(\boldsymbol{M};K) \to \mathscr{A}^\bullet(\boldsymbol{N};\psi(K))$ *(this applies in particular to the (time-)orientation preserving isometric isomorphisms of* $\boldsymbol{M}$*).*

Proof: (a) Immediate from the definition. (b) These results follow from (a) because $J_{\boldsymbol{M}}(K) = J_{\boldsymbol{M}}(K^{\perp\perp})$ for compact K (see Lem. A.10(ii)) and the obvious inclusions $K_1 \cap K_2 \subset K_i \subset K_1 \cup K_2$. (c) As ψ is an isomorphism, $\psi(K)^\perp = \psi(K^\perp)$. Thus the pushforward ψ_* restricts to an isomorphism between $H(\boldsymbol{M};K^\perp)$ and $H(\boldsymbol{N};\psi(K)^\perp)$, with inverse given by the pullback ψ^*. Hence for all $\boldsymbol{h} \in H(\boldsymbol{N};\psi(K))$,

$$\mathrm{rce}_{\boldsymbol{N}}[\boldsymbol{h}] \circ \mathscr{A}(\psi) \circ \alpha^\bullet_{\boldsymbol{M};K} = \mathscr{A}(\psi) \circ \mathrm{rce}_{\boldsymbol{M}}[\psi^*\boldsymbol{h}] \circ \alpha^\bullet_{\boldsymbol{M};K} = \mathscr{A}(\psi) \circ \alpha^\bullet_{\boldsymbol{M};K}$$

by the defining property of $\alpha^\bullet_{\boldsymbol{M};K}$; it follows that $\mathscr{A}(\psi) \circ \alpha^\bullet_{\boldsymbol{M};K} = \alpha^\bullet_{\boldsymbol{N};\psi(K)} \circ \beta$ for some β (depending on ψ and K). Applying the same argument to ψ^{-1}, it follows easily that β is an isomorphism. $\square$

These results immediately induce a number of analogous properties of the $\mathscr{A}^{\mathrm{dyn}}(\boldsymbol{M};O)$, in Theorem 5.4 below. First, we give a useful simplifying observation.

Lemma 5.3 *Given any* $O \in \mathscr{O}(\boldsymbol{M})$*, we have*

$$\mathscr{A}^{\mathrm{dyn}}(\boldsymbol{M};O) \cong \bigvee_{K \in \mathscr{K}_b(\boldsymbol{M};O)} \mathscr{A}^\bullet(\boldsymbol{M};K),$$

where $\mathscr{K}_b(\boldsymbol{M};O)$ *is the set of those* $K \in \mathscr{K}(\boldsymbol{M};O)$ *obtained as the closure of a base of a multi-diamond, with* $\mathscr{K}_b(\boldsymbol{M};\emptyset) = \{\emptyset\}$ *by convention. If, in fact,* O *is a multi-diamond, then*

$$\mathscr{A}^{\mathrm{dyn}}(\boldsymbol{M};O) \cong \bigvee_{K \subset\subset B} \mathscr{A}^\bullet(\boldsymbol{M};K),$$

where B *is any base of* O *and the union is taken over all compact subsets of* B*.*

Proof: If O is empty, the first statement holds trivially because $\mathscr{K}_b(\boldsymbol{M};\emptyset) = \mathscr{K}(\boldsymbol{M};\emptyset) = \{\emptyset\}$; as $\emptyset$ is not a multi-diamond the second statement is irrelevant. Accordingly, now assume that O is nonempty and let $K \in \mathscr{K}(\boldsymbol{M};O)$. Then there is a multi-diamond with base $B \subset O$ such that $K \subset D_{\boldsymbol{M}}(B)$. By Lemma A.14, there exists a compact set $\tilde{K} \subset B$ with $K \subset \tilde{K}^{\perp\perp}$; it is clear that $\tilde{K} \subset \mathscr{K}(\boldsymbol{M};O)$. Hence $\mathscr{A}^\bullet(\boldsymbol{M};K) \subset \mathscr{A}^\bullet(\boldsymbol{M};\tilde{K}^{\perp\perp}) \cong \mathscr{A}^\bullet(\boldsymbol{M};\tilde{K})$ by parts (a) and (c) of Thm. 5.2. In the case of a general nonempty $O \in \mathscr{O}(\boldsymbol{M})$ we deduce that the defining union of $\mathscr{A}^{\mathrm{dyn}}(\boldsymbol{M};O)$ may be taken over $K \in \mathscr{K}_b(\boldsymbol{M};O)$ [see Lemma B.2 for a proof in the abstract setting]; in the case where O is a multi-diamond with B as a base, we may evidently require that each $\tilde{K}$ be a subset of B, obtaining the second refinement (every compact subset of B is clearly a member of $\mathscr{K}(\boldsymbol{M};O)$). $\square$

We expect that stronger causality results than (c) below can be obtained along similar lines.

Theorem 5.4 *(a) If $O_1, O_2 \in \mathscr{O}(\boldsymbol{M})$ and $O_1 \subset O_2$ then $\mathscr{A}^{\rm dyn}(\boldsymbol{M}; O_1) \subset \mathscr{A}^{\rm dyn}(\boldsymbol{M}; O_2)$. In consequence, we also have, for arbitrary $O_1, O_2 \in \mathscr{O}(\boldsymbol{M})$,*

$$\begin{aligned} \mathscr{A}^{\rm dyn}(\boldsymbol{M}; O_1) \vee \mathscr{A}^{\rm dyn}(\boldsymbol{M}; O_2) &\subset \mathscr{A}^{\rm dyn}(\boldsymbol{M}; O_1 \cup O_2) \\ \mathscr{A}^{\rm dyn}(\boldsymbol{M}; O_1 \cap O_2) &\subset \mathscr{A}^{\rm dyn}(\boldsymbol{M}; O_1) \wedge \mathscr{A}^{\rm dyn}(\boldsymbol{M}; O_2) \end{aligned}$$

and $\mathscr{A}^\bullet(\boldsymbol{M}; \emptyset) \cong \mathscr{A}^{\rm dyn}(\boldsymbol{M}; \emptyset) \subset \mathscr{A}^{\rm dyn}(\boldsymbol{M}; O)$ for all $O \in \mathscr{O}(\boldsymbol{M})$.
(b) If $\psi : \boldsymbol{M} \to \boldsymbol{N}$ is an isomorphism then $\mathscr{A}(\psi)$ restricts to an isomorphism $\mathscr{A}^{\rm dyn}(\boldsymbol{M}; O) \to \mathscr{A}^{\rm dyn}(\boldsymbol{N}; \psi(O))$ for each $O \in \mathscr{O}(\boldsymbol{M})$. (In particular, this applies to automorphisms $\psi \in \mathrm{Aut}(\boldsymbol{M})$.)
(c) If $O \in \mathscr{O}(\boldsymbol{M})$ and O'' is a multi-diamond with a base contained in O, then

$$\mathscr{A}^{\rm dyn}(\boldsymbol{M}; O'') \cong \mathscr{A}^{\rm dyn}(\boldsymbol{M}; O).$$

Proof: (a) is obvious because $\mathscr{K}(\boldsymbol{M}; O_1) \subset \mathscr{K}(\boldsymbol{M}; O_2)$. For (b), we use Thm. 5.2(c) and the obvious fact that the unions of isomorphic subobjects of isomorphic objects are isomorphic. Turning to (c), we may suppose that $O'' = D_{\boldsymbol{M}}(B)$, where $B \subset O$ is a base of O''. By Lemma 5.3 we then have

$$\mathscr{A}^{\rm dyn}(\boldsymbol{M}; O'') \cong \bigvee_{K \subset\subset B} \mathscr{A}^\bullet(\boldsymbol{M}; K) \subset \mathscr{A}^{\rm dyn}(\boldsymbol{M}; O) \subset \mathscr{A}^{\rm dyn}(\boldsymbol{M}; O''),$$

where we have also used part (a). □

Further light on the relationship between the two species of dynamical net is shed by the next result. We will need the following definition.

Definition 5.5 *A compact set $K \subset \boldsymbol{M}$ will be called* outer regular *if there exist relatively compact nonempty $O_n \in \mathscr{O}(\boldsymbol{M})$ ($n \in \mathbb{N}$) with $\mathrm{cl}(O_{n+1}) \subset O_n$ and $K \in \mathscr{K}(\boldsymbol{M}; O_n)$ for all n, such that $K = \bigcap_n O_n$. [Note that this excludes the empty set from being outer regular.] Any such sequence O_n will be called an* outer approximation *to K. The set of outer regular compact subsets of any nonempty $O \in \mathscr{O}(\boldsymbol{M})$ will be denoted $\mathscr{K}^{o.r.}(\boldsymbol{M}; O)$. If $K \in \mathscr{K}^{o.r.}(\boldsymbol{M}; O)$ has an outer approximating sequence $O_n \in \mathscr{O}_0(\boldsymbol{M})$, we write $K \in \mathscr{K}_0^{o.r.}(\boldsymbol{M}; O)$.*

Note that $\mathscr{K}_b(\boldsymbol{M}; O) \subset \mathscr{K}^{o.r.}(\boldsymbol{M}; O)$. We write $\mathscr{K}^{o.r.}(\boldsymbol{M})$ for $\mathscr{K}^{o.r.}(\boldsymbol{M}; \mathcal{M})$.

Theorem 5.6 *(a) For all $O \in \mathscr{O}(\boldsymbol{M})$ and $\boldsymbol{h} \in H(\boldsymbol{M}; O')$ we have $\mathrm{rce}_{\boldsymbol{M}}[\boldsymbol{h}] \circ \alpha^{\rm dyn}_{\boldsymbol{M};O} = \alpha^{\rm dyn}_{\boldsymbol{M};O}$.*
(b) If $O \in \mathscr{O}(\boldsymbol{M})$ is relatively compact, then

$$\mathscr{A}^{\rm dyn}(\boldsymbol{M}; O) \subset \mathscr{A}^\bullet(\boldsymbol{M}; \mathrm{cl}(O)).$$

(c) If $K \in \mathscr{K}^{o.r.}(\boldsymbol{M})$ has outer approximating sequence O_n, then

$$\mathscr{A}^\bullet(\boldsymbol{M}; K) \cong \bigwedge_{n \in \mathbb{N}} \mathscr{A}^{\rm dyn}(\boldsymbol{M}; O_n). \tag{5.4}$$

Proof: (a) If $K \in \mathscr{K}(\boldsymbol{M};O)$, then $K^\perp \supset O'$ and hence $\boldsymbol{h} \in H(\boldsymbol{M};K^\perp)$. Thus $\mathrm{rce}_{\boldsymbol{M}}[\boldsymbol{h}] \circ \alpha^\bullet_{\boldsymbol{M};K} = \alpha^\bullet_{\boldsymbol{M};K}$ for all such K. The same then holds for $\alpha^{\mathrm{dyn}}_{\boldsymbol{M};O}$ due to Eq. (5.3). (For completeness, a proof is given in Lem. B.3 of Appendix B.)
(b) Lemma A.12 entails that $O' = (\mathrm{cl}(O))^\perp$. Using (a), we deduce that $\mathrm{rce}_{\boldsymbol{M}}[\boldsymbol{h}] \circ \alpha^{\mathrm{dyn}}_{\boldsymbol{M};O} = \alpha^{\mathrm{dyn}}_{\boldsymbol{M};O}$ for all $\boldsymbol{h} \in H(\boldsymbol{M};\mathrm{cl}(O)^\perp)$ and hence $\alpha^{\mathrm{dyn}}_{\boldsymbol{M};O} \le \alpha^\bullet_{\boldsymbol{M};\mathrm{cl}(O)}$, establishing the required inclusion.
(c) As $K \in \mathscr{K}(\boldsymbol{M};O_n)$ for each $n \in \mathbb{N}$ the right-hand side of Eq. (5.4) clearly contains the left-hand side. On the other hand, Lemma A.11(ii) entails that

$$K^\perp = \bigcup_{n\in\mathbb{N}} O'_n$$

so for any $\boldsymbol{h} \in H(\boldsymbol{M};K^\perp)$, the compact set $\mathrm{supp}\,\boldsymbol{h} \subset K^\perp$ is covered by finitely many of the open sets O'_n and hence (as $O'_n \subset O'_{n+1}$ for each n) is contained in some O'_{n_0}. It follows that $\boldsymbol{h} \in H(\boldsymbol{M};O'_{n_0})$, so $\mathrm{rce}_{\boldsymbol{M}}[\boldsymbol{h}]$ acts trivially on $\mathscr{A}^{\mathrm{dyn}}(\boldsymbol{M};O_{n_0})$ and therefore on the intersection in Eq. (5.4). Accordingly, the right-hand side is contained in $\mathscr{A}^\bullet(\boldsymbol{M};K)$. $\square$

We remark that this result also gives $\mathscr{A}^\bullet(\boldsymbol{M};K)$ for sets K such that $K^{\perp\perp}$ is outer regular, by virtue of Thm 5.2(b).

As an example of the various relationships developed above, we note that if p, q are distinct timelike separated points with q to the future of p, then

$$\mathscr{A}^\bullet(\boldsymbol{M};\{p,q\}) \cong \mathscr{A}^\bullet(\boldsymbol{M};\{p,q\}^{\perp\perp}) \cong \mathscr{A}^\bullet(\boldsymbol{M};J^+_{\boldsymbol{M}}(p)\cap J^-_{\boldsymbol{M}}(q)) \supset \mathscr{A}^{\mathrm{dyn}}(\boldsymbol{M};I^+_{\boldsymbol{M}}(p)\cap I^-_{\boldsymbol{M}}(q)).$$

Moreover, if $p_n \to p$ in $I^-_{\boldsymbol{M}}(p)$, and $q_n \to q$ in $I^+_{\boldsymbol{M}}(q)$ then

$$\mathscr{A}^\bullet(\boldsymbol{M};\{p,q\}) \cong \bigwedge_n \mathscr{A}^{\mathrm{dyn}}(\boldsymbol{M};I^+_{\boldsymbol{M}}(p_n)\cap I^-_{\boldsymbol{M}}(q_n)).$$

In addition if $K \in \mathscr{K}(\boldsymbol{M})$ is the closure of a Cauchy multi-ball, then we may choose a sequence of Cauchy multi-balls B_k such that $\mathrm{cl}(B_{k+1}) \subset B_k$ and $\bigcap_k B_k = K$. Choose a strictly decreasing sequence (ϵ_k) with $\epsilon_k \to 0$ such that $O_k = D_{\boldsymbol{M}}(B_k) \cap \mathcal{T}^{-1}(-\epsilon_k, \epsilon_k)$ belongs to $\mathscr{O}(\boldsymbol{M})$ for each k, where $\mathcal{T}$ is a Cauchy temporal function [4] such that $\mathcal{T}^{-1}(0)$ contains all the B_k. Then $K = \bigcap_k O_k$ and so

$$\mathscr{A}^\bullet(\boldsymbol{M};K) \cong \bigwedge_k \mathscr{A}^{\mathrm{dyn}}(\boldsymbol{M};O_k).$$

Finally, let us compute the dynamical nets of ordinary diagonal models φ_Δ.

Theorem 5.7 *For any ordinary diagonal theory φ_Δ, we have*

$$\varphi^\bullet_\Delta(\boldsymbol{M};K) = \varphi(\boldsymbol{M})^\bullet(\boldsymbol{M};K) \tag{5.5}$$

for all compact $K \subset \boldsymbol{M}$, and

$$\varphi^{\mathrm{dyn}}_\Delta(\boldsymbol{M};O) = \varphi(\boldsymbol{M})^{\mathrm{dyn}}(\boldsymbol{M};O) \tag{5.6}$$

for all $O \in \mathscr{O}(\boldsymbol{M})$.

Proof: The first statement is an immediate consequence from Prop. 4.5 as $\mathrm{rce}_{\boldsymbol{M}}^{(\varphi_\Delta)}[\boldsymbol{h}] = \mathrm{rce}_{\boldsymbol{M}}^{(\varphi(\boldsymbol{M}))}[\boldsymbol{h}]$; the second follows immediately. □

Thus the ordinary diagonal theories provide examples in which it is the dynamical net, rather than the kinematic net, that appears to have the 'right' notion of the local observables on any given spacetime. (As we have no examples of extraordinary diagonal theories, it is less clear what should be expected in that case.)

6 Dynamical locality

6.1 Definition and main properties

The kinematical and dynamical nets give two isotonous nets on each spacetime; the diagonal theories show that they are not always equal. In general, their relationship is given as follows.

Proposition 6.1 *Let $\mathscr{A} \in \mathsf{LCT}$ (resp., LCT_0). Suppose $O \in \mathscr{O}(\boldsymbol{M})$ (resp., $\mathscr{O}_0(\boldsymbol{M})$) is nonempty, and that $O \subset K \in \mathscr{K}(\boldsymbol{M}; \tilde{O})$ for some $\tilde{O} \in \mathscr{O}(\boldsymbol{M})$. Then*

$$\mathscr{A}^{\mathrm{kin}}(\boldsymbol{M}; O) \subset \mathscr{A}^{\bullet}(\boldsymbol{M}; K) \subset \mathscr{A}^{\mathrm{dyn}}(\boldsymbol{M}; \tilde{O}).$$

Proof: By Prop. 3.5, we have $\mathrm{rce}_{\boldsymbol{M}}[\boldsymbol{h}] \circ \alpha^{\mathrm{kin}}_{\boldsymbol{M};O} = \alpha^{\mathrm{kin}}_{\boldsymbol{M};O}$ for all $\boldsymbol{h} \in H(\boldsymbol{M}; K^{\perp})$, so $\alpha^{\mathrm{kin}}_{\boldsymbol{M};O} = \alpha^{\bullet}_{\boldsymbol{M};K} \circ \beta$ for some β and the first inclusion is proved. The second follows immediately as $\mathscr{A}^{\bullet}(\boldsymbol{M}; K)$ is one of the generating algebras for $\mathscr{A}^{\mathrm{dyn}}(\boldsymbol{M}; \tilde{O})$. □

A clear case of interest is that in which these two nets actually coincide; in view of Prop. 6.1 this is a maximality condition on the kinematic net. It requires, roughly, that every observable invariant under changes of metric in the causal complement of O is localised in O.

Definition 6.2 *A theory $\mathscr{A} \in \mathsf{LCT}$ (resp., LCT_0) obeys* dynamical locality *if it obeys the timeslice property and, additionally, for each $\boldsymbol{M} \in \mathsf{Loc}$ (resp., Loc_0) and all nonempty $O \in \mathscr{O}(\boldsymbol{M})$ (resp., $\mathscr{O}_0(\boldsymbol{M})$) we have $\mathscr{A}^{\mathrm{kin}}(\boldsymbol{M}; O) \cong \mathscr{A}^{\mathrm{dyn}}(\boldsymbol{M}; O)$, i.e., more abstractly,*

$$\alpha^{\mathrm{kin}}_{\boldsymbol{M};O} \cong \alpha^{\mathrm{dyn}}_{\boldsymbol{M};O}.$$

In view of Lem. 3.1, the dynamical locality condition may also be written in the form

$$\mathscr{A}(\psi) \cong \alpha^{\mathrm{dyn}}_{\boldsymbol{N};\psi(\boldsymbol{M})} \cong \bigvee_{K \in \mathscr{K}(\boldsymbol{N};\psi(\boldsymbol{M}))} \alpha^{\bullet}_{\boldsymbol{N};K}$$

for all $\psi : \boldsymbol{M} \to \boldsymbol{N}$.

An immediate example is furnished by the initial theory $\mathscr{I}$, because all subobjects of an initial object are isomorphic. More physically interesting theories will be considered in [28]. In the remainder of this section we explore various general features of dynamically local theories without restricting Phys; later, in section 6.3, we will consider applications to quantum field theory by specifying that Phys should be Alg or C^*-Alg.

Additivity Dynamical locality imposes a form of additivity on the theory.

Theorem 6.3 *Suppose $\mathscr{A} \in \mathsf{LCT}$ (resp., LCT_0) is dynamically local in LCT (resp., LCT_0). (a) For any $\boldsymbol{M} \in \mathsf{Loc}$ (resp., Loc_0), the maps $\bigvee_{K\in\mathscr{K}(\boldsymbol{M})} \alpha^\bullet_{\boldsymbol{M};K}$ and $\bigvee_{K\in\mathscr{K}_b(\boldsymbol{M})} \alpha^\bullet_{\boldsymbol{M};K}$ are isomorphisms, i.e.,*

$$\mathscr{A}(\boldsymbol{M}) \cong \bigvee_{K\in\mathscr{K}(\boldsymbol{M})} \mathscr{A}^\bullet(\boldsymbol{M};K) \cong \bigvee_{K\in\mathscr{K}_b(\boldsymbol{M})} \mathscr{A}^\bullet(\boldsymbol{M};K).$$

(b) Suppose $\widetilde{\mathscr{O}}$ is a subset of $\mathscr{O}(\boldsymbol{M})$ such that every $K \in \mathscr{K}_b(\boldsymbol{M})$ is contained in some $O \in \widetilde{\mathscr{O}}$. Then $\bigvee_{O\in\widetilde{\mathscr{O}}} \alpha^{\rm dyn}_{\boldsymbol{M};O}$ is an isomorphism, i.e.,

$$\mathscr{A}(\boldsymbol{M}) \cong \bigvee_{O\in\widetilde{\mathscr{O}}} \mathscr{A}^{\rm dyn}(\boldsymbol{M};O) \cong \bigvee_{O\in\widetilde{\mathscr{O}}} \mathscr{A}^{\rm kin}(\boldsymbol{M};O).$$

Remark: In particular, by definition of $\mathscr{K}_b(\boldsymbol{M})$, part (b) applies when the $\widetilde{\mathscr{O}}$ consists of the truncated multi-diamonds of $\boldsymbol{M}$.
Proof: (a) First observe that $\alpha^{\rm dyn}_{\boldsymbol{M};\mathcal{M}} \cong \alpha^{\rm kin}_{\boldsymbol{M};\mathcal{M}} \cong {\rm id}_{\mathscr{A}(\boldsymbol{M})}$ (by Lemma 3.1). Thus $\alpha^{\rm dyn}_{\boldsymbol{M};\mathcal{M}}$ is an isomorphism. The statement follows from the definition of $\alpha^{\rm dyn}_{\boldsymbol{M};\mathcal{M}}$ and Lemma 5.3.

(b) For each $K \in \mathscr{K}(\boldsymbol{M})$ choose a $O_K \in \widetilde{\mathscr{O}}$ with $K \subset O_K$, whereupon there is a factorization $\alpha^\bullet_{\boldsymbol{M};K} = \alpha^{\rm dyn}_{\boldsymbol{M};O_K} \circ \alpha_{\boldsymbol{M};O_K;O}$ for each such K. By Lemma B.2,

$$\bigvee_{K\in\mathscr{K}_b(\boldsymbol{M})} \alpha^\bullet_{\boldsymbol{M};K} \le \bigvee_{O\in\widetilde{\mathscr{O}}} \alpha^{\rm dyn}_{\boldsymbol{M};O}.$$

As the left-hand side is an isomorphism, the monic property of $\bigvee_{O\in\widetilde{\mathscr{O}}} \alpha^{\rm dyn}_{\boldsymbol{M};O}$ implies that it is an isomorphism. The remaining statements are immediate. □

Covariance Theorem 5.2(b) and Theorem 5.4(b) provide rather weaker forms of covariance than the relation Eq. (3.3) that holds for the kinematic net. Dynamical locality provides the missing ingredient, provided the class of compact indexing regions is restricted slightly.

Theorem 6.4 *Suppose $\mathscr{A} \in \mathsf{LCT}$ (resp., LCT_0) is dynamically local in LCT (resp., LCT_0) and let $\psi : \boldsymbol{M} \to \boldsymbol{N}$ in Loc (resp., LCT_0). Then for all nonempty $\mathscr{O} \in \mathscr{O}(\boldsymbol{M})$ and $K \in \mathscr{K}^{o.r.}(\boldsymbol{M})$ (resp., $\mathscr{O}_0(\boldsymbol{M})$, $K \in \mathscr{K}_0^{o.r.}(\boldsymbol{M})$), we have*

$$\alpha^{\rm dyn}_{\boldsymbol{N};\psi(O)} \cong \mathscr{A}(\psi) \circ \alpha^{\rm dyn}_{\boldsymbol{M};O} \qquad \textit{and} \qquad \alpha^\bullet_{\boldsymbol{N};\psi(K)} \cong \mathscr{A}(\psi) \circ \alpha^\bullet_{\boldsymbol{M};K}.$$

The second formula holds also for compact K such that $K^{\perp\perp}$ is outer regular.

Proof: The first statement follows immediately from dynamical locality and the covariance of the kinematic net of Eq. (3.3), by the calculation

$$\alpha^{\rm dyn}_{\boldsymbol{N};\psi(O)} \cong \alpha^{\rm kin}_{\boldsymbol{N};\psi(O)} = \mathscr{A}(\psi) \circ \alpha^{\rm kin}_{\boldsymbol{M};O} \cong \mathscr{A}(\psi) \circ \alpha^{\rm dyn}_{\boldsymbol{M};O}.$$

For the second, we claim that if O_n is outer approximating to K in $\boldsymbol{M}$, then $\psi(O_n)$ is outer approximating to $\psi(K)$ in $\boldsymbol{N}$. We use the fact that ψ maps diamonds and their bases in $\boldsymbol{M}$ to diamonds and their bases in $\boldsymbol{N}$; this is otherwise straightforward. Using this observation and the first part of the result, we calculate

$$\alpha^{\bullet}_{\boldsymbol{N};\psi(K)} \cong \bigwedge_n \alpha^{\rm dyn}_{\boldsymbol{N};\psi(O_n)} \cong \bigwedge_n \mathscr{A}(\psi)\circ\alpha^{\rm dyn}_{\boldsymbol{M};O_n} \cong \mathscr{A}(\psi)\circ\bigwedge_n \alpha^{\rm dyn}_{\boldsymbol{M};O_n} \cong \mathscr{A}(\psi)\circ\alpha^{\bullet}_{\boldsymbol{M};K},$$

in conjunction with Thm. 5.2(b) and Lem. B.1. Finally, if $K^{\perp\perp} \in \mathscr{K}^{o.r.}(\boldsymbol{M})$ (resp., $\mathscr{K}_0^{o.r.}(\boldsymbol{M})$), we calculate

$$\alpha^{\bullet}_{\boldsymbol{N};\psi(K)} \cong \alpha^{\bullet}_{\boldsymbol{N};\psi(K)^{\perp\perp}} = \alpha^{\bullet}_{\boldsymbol{N};\psi(K^{\perp\perp})} \cong \mathscr{A}(\psi)\circ\alpha^{\bullet}_{\boldsymbol{M};K^{\perp\perp}} \cong \mathscr{A}(\psi)\circ\alpha^{\bullet}_{\boldsymbol{M};K}$$

using the previous result, Thm. 5.2(b), and the identity $\psi(K^{\perp\perp}) = \psi(K)^{\perp\perp}$ proved in Lemma A.15. □

Extended locality In Minkowski space algebraic QFT, extended locality [49, 38] is the condition that local algebras of spacelike separated regions should intersect only on multiples of the identity. Here, we will give a necessary and sufficient condition for a version of extended locality in general locally covariant physical theories subject to dynamical locality.

Theorem 6.5 *Suppose that $\mathscr{A} \in \mathsf{LCT}$ (resp., LCT_0) is dynamically local. Then the following are equivalent:*

1. *$\mathscr{A}$ obeys extended locality, in the sense that $\alpha^{\rm kin}_{\boldsymbol{M};O_1}\wedge\alpha^{\rm kin}_{\boldsymbol{M};O_2}$ is trivial for all causally disjoint nonempty $O_i \in \mathscr{O}(\boldsymbol{M})$[17] (resp., $\mathscr{O}_0(\boldsymbol{M})$) for arbitrary $\boldsymbol{M} \in \mathsf{Loc}$ (resp., Loc_0);*
2. *$\alpha^{\bullet}_{\boldsymbol{M};\emptyset}$ (or equivalently $\alpha^{\rm dyn}_{\boldsymbol{M};\emptyset}$) is trivial, i.e., equivalent to $\mathcal{I}_{\mathscr{A}(\boldsymbol{M})}$ for every $\boldsymbol{M} \in \mathsf{Loc}$ (resp., Loc_0).*

Proof: (1) $\Longrightarrow$ (2): take any two nonempty causally disjoint $O_i \in \mathscr{O}_0(\boldsymbol{M})$. We then have, using Thm. 5.4(a),

$$\alpha^{\bullet}_{\boldsymbol{M};\emptyset} \cong \alpha^{\rm dyn}_{\boldsymbol{M};\emptyset} = \alpha^{\rm dyn}_{\boldsymbol{M};O_1\cap O_2} \le \alpha^{\rm dyn}_{\boldsymbol{M};O_1}\wedge\alpha^{\rm dyn}_{\boldsymbol{M};O_2} \cong \alpha^{\rm kin}_{\boldsymbol{M};O_1}\wedge\alpha^{\rm kin}_{\boldsymbol{M};O_2} \cong \mathcal{I}_{\mathscr{A}(\boldsymbol{M})}.$$

(2) $\Longrightarrow$ (1): On the other hand, let $\alpha \cong \alpha^{\rm kin}_{\boldsymbol{M};O_1}\wedge\alpha^{\rm kin}_{\boldsymbol{M};O_2}$ for causally disjoint nonempty $O_i \in \mathscr{O}_0(\boldsymbol{M})$. Then $\alpha = \alpha^{\rm kin}_{\boldsymbol{M};O_i}\circ\alpha_i$ for some α_i; we will show that the α_i are trivial, which implies triviality of α. To this end, let $\boldsymbol{h} \in H(\boldsymbol{M}|_{O_1})$ be arbitrary and observe that

$$\begin{aligned}\mathscr{A}(\iota_{\boldsymbol{M};O_1})\circ{\rm rce}_{\boldsymbol{M}|_{O_1}}[\boldsymbol{h}]\circ\alpha_1 &= {\rm rce}_{\boldsymbol{M}}[\iota_{\boldsymbol{M};O_1*}\boldsymbol{h}]\circ\mathscr{A}(\iota_{\boldsymbol{M};O_1})\circ\alpha_1 = {\rm rce}_{\boldsymbol{M}}[\iota_{\boldsymbol{M};O_1*}\boldsymbol{h}]\circ\mathscr{A}(\iota_{\boldsymbol{M};O_2})\circ\alpha_2\\ &= \mathscr{A}(\iota_{\boldsymbol{M};O_2})\circ\alpha_2 = \mathscr{A}(\iota_{\boldsymbol{M};O_1})\circ\alpha_1,\end{aligned}$$

[17]That is, $O_1 \subset O_2^{\perp}$ and $O_2 \subset O_1^{\perp}$, from which it follows that $O_1 \subset O_2'$ and $O_2 \subset O_1'$.

where we have used the causal separation of the O_i and Prop. 3.5. Cancelling the monic $\mathscr{A}(\iota_{\boldsymbol{M};O_1})$, we have $\mathrm{rce}_{\boldsymbol{M}|_{O_1}}[\boldsymbol{h}] \circ \alpha_1 = \alpha_1$ for all $\boldsymbol{h} \in H(\boldsymbol{M}|_{O_1})$. Hence $\alpha_1 \le \alpha^\bullet_{\boldsymbol{M}|_O;\emptyset} \cong \mathcal{I}_{\mathscr{A}(\boldsymbol{M}|_{O_1})}$ and is therefore trivial. □

The subobject $\alpha^\bullet_{\boldsymbol{M};\emptyset}$ represents those elements of the theory that are invariant with respect to arbitrary perturbations of the metric, and therefore do not couple to gravity. Under many circumstances one would want this to be trivial, i.e., that $\alpha^\bullet_{\boldsymbol{M};\emptyset} \cong \mathcal{I}_{\boldsymbol{M}}$ for all spacetimes $\boldsymbol{M}$. As we will see, this requirement is not always satisfied – indeed, it is not satisfied for the theory of the free massless minimally coupled scalar field in spacetimes of compact spatial section. However, it can be derived from other reasonable conditions on theories in LCT as will be discussed elsewhere.

6.2 The SPASs property

The pathological theories constructed in Sect. 4 had the property that there are natural transformations between them such that some, but not all, of their components are isomorphisms. In this section we prove that this cannot occur if we restrict to dynamically local theories. Throughout this section, $\mathscr{A}$ and $\mathscr{B}$ are fixed theories in either LCT or LCT_0 obeying the timeslice property.

The following preparatory lemmas are elementary, but crucial; we give proofs for completeness.

Lemma 6.6 *Let $\boldsymbol{M}$ be an arbitrary spacetime. Suppose there is a morphism $\zeta_{\boldsymbol{M}} : \mathscr{B}(\boldsymbol{M}) \to \mathscr{A}(\boldsymbol{M})$ [not necessarily a component of a natural transformation] such that*

$$\mathrm{rce}^{(\mathscr{A})}_{\boldsymbol{M}}[\boldsymbol{h}] \circ \zeta_{\boldsymbol{M}} = \zeta_{\boldsymbol{M}} \circ \mathrm{rce}^{(\mathscr{B})}_{\boldsymbol{M}}[\boldsymbol{h}] \tag{6.1}$$

for all $\boldsymbol{h} \in H(\boldsymbol{M}; K^\perp)$. Then there are unique morphisms

$$\zeta^\bullet_{\boldsymbol{M};K} : \mathscr{B}^\bullet(\boldsymbol{M}; K) \to \mathscr{A}^\bullet(\boldsymbol{M}; K)$$
$$\zeta^{\mathrm{dyn}}_{\boldsymbol{M};O} : \mathscr{B}^{\mathrm{dyn}}(\boldsymbol{M}; O) \to \mathscr{A}^{\mathrm{dyn}}(\boldsymbol{M}; O)$$

such that

$$\alpha^\bullet_{\boldsymbol{M};K} \circ \zeta^\bullet_{\boldsymbol{M};K} = \zeta_{\boldsymbol{M}} \circ \beta^\bullet_{\boldsymbol{M};K} \tag{6.2}$$
$$\alpha^{\mathrm{dyn}}_{\boldsymbol{M};O} \circ \zeta^{\mathrm{dyn}}_{\boldsymbol{M};O} = \zeta_{\boldsymbol{M}} \circ \beta^{\mathrm{dyn}}_{\boldsymbol{M};O}, \tag{6.3}$$

where we use $\beta^\bullet_{\boldsymbol{M};K}$ and $\beta^{\mathrm{int}}_{\boldsymbol{M};O}$ for the inclusion morphisms of $\mathscr{B}^\bullet(\boldsymbol{M}; K)$ and $\mathscr{B}^{\mathrm{dyn}}(\boldsymbol{M}; O)$ in $\mathscr{B}(\boldsymbol{M})$. Thus $\zeta^\bullet_{\boldsymbol{M};K}$ and $\zeta^{\mathrm{dyn}}_{\boldsymbol{M};O}$ are restrictions of $\zeta_{\boldsymbol{M}}$. Moreover, if $\zeta_{\boldsymbol{M}}$ is an isomorphism, so are $\zeta^\bullet_{\boldsymbol{M};K}$ and $\zeta^{\mathrm{dyn}}_{\boldsymbol{M};O}$. In particular, these conclusions hold if $\zeta_{\boldsymbol{M}}$ is a component of a natural transformation $\zeta : \mathscr{B} \dot{\to} \mathscr{A}$.

Proof: As $\alpha^\bullet_{\boldsymbol{M};K}$ and $\alpha^{\mathrm{dyn}}_{\boldsymbol{M};O}$ are monic, uniqueness is automatic and one need only demonstrate existence. First, by Eq. (6.1) and the defining property of $\beta^\bullet_{\boldsymbol{M};K}$,

$$\mathrm{rce}^{(\mathscr{A})}_{\boldsymbol{M}}[\boldsymbol{h}] \circ \zeta_{\boldsymbol{M}} \circ \beta^\bullet_{\boldsymbol{M};K} = \zeta_{\boldsymbol{M}} \circ \mathrm{rce}^{(\mathscr{B})}_{\boldsymbol{M}}[\boldsymbol{h}] \circ \beta^\bullet_{\boldsymbol{M};K} = \zeta_{\boldsymbol{M}} \circ \beta^\bullet_{\boldsymbol{M};K}$$

for all $\boldsymbol{h} \in H(\boldsymbol{M}; K^\perp)$. Hence $\zeta_{\boldsymbol{M}} \circ \beta^\bullet_{\boldsymbol{M};K}$ shares the defining property of $\alpha^\bullet_{\boldsymbol{M};K}$ and we deduce the existence of unique $\zeta^\bullet_{\boldsymbol{M};K} : \mathscr{B}^\bullet(\boldsymbol{M}; K) \to \mathscr{A}^\bullet(\boldsymbol{M}; K)$ such that Eq. (6.2) holds.

Second, for each $\mathscr{K}(\boldsymbol{M}) \ni K \subset O$, the outer portion of the diagram

$$\begin{tikzcd}
\mathscr{B}^\bullet(\boldsymbol{M}; K) \arrow[r, "\beta_{\boldsymbol{M};O;K}"] \arrow[d, "\zeta^\bullet_{\boldsymbol{M};K}"'] & \mathscr{B}^{\mathrm{dyn}}(\boldsymbol{M}; O) \arrow[r, "\beta^{\mathrm{dyn}}_{\boldsymbol{M};O}"] \arrow[d, dotted, "\zeta^{\mathrm{dyn}}_{\boldsymbol{M};O}"'] & \mathscr{B}(\boldsymbol{M}) \arrow[d, "\zeta_{\boldsymbol{M}}"'] \\
\mathscr{A}^\bullet(\boldsymbol{M}; K) \arrow[r, "\alpha_{\boldsymbol{M};O;K}"] & \mathscr{A}^{\mathrm{dyn}}(\boldsymbol{M}; O) \arrow[r, "\alpha^{\mathrm{dyn}}_{\boldsymbol{M};O}"] & \mathscr{A}(\boldsymbol{M})
\end{tikzcd}$$

now commutes, thus inducing a unique $\zeta^{\mathrm{dyn}}_{\boldsymbol{M};O} : \mathscr{B}^{\mathrm{dyn}}(\boldsymbol{M}; O) \to \mathscr{A}^{\mathrm{int}}(\boldsymbol{M}; O)$ such that all the diagrams commute in full; in particular, we have the required property Eq. (6.3).

Thirdly, if $\zeta_{\boldsymbol{M}}$ is an isomorphism, Eq. (6.1) holds with $\zeta_{\boldsymbol{M}}$ replaced by $\zeta^{-1}_{\boldsymbol{M}}$ and $\mathscr{A}$ and $\mathscr{B}$ interchanged. Thus there are unique morphisms $(\zeta^{-1}_{\boldsymbol{M}})^\bullet_{;K}$ and $(\zeta^{-1}_{\boldsymbol{M}})^{\mathrm{int}}_{;O}$ such that

$$\begin{aligned}
\beta^\bullet_{\boldsymbol{M};K} \circ (\zeta^{-1}_{\boldsymbol{M}})^\bullet_{;K} &= \zeta^{-1}_{\boldsymbol{M}} \circ \alpha^\bullet_{\boldsymbol{M};K} \\
\beta^{\mathrm{dyn}}_{\boldsymbol{M};O} \circ (\zeta^{-1}_{\boldsymbol{M}})^{\mathrm{dyn}}_{;O} &= \zeta^{-1}_{\boldsymbol{M}} \circ \alpha^{\mathrm{dyn}}_{\boldsymbol{M};O}.
\end{aligned}$$

Combining with Eqs. (6.2) and (6.3) and using the facts that $\alpha^\bullet_{\boldsymbol{M};K}, \beta^\bullet_{\boldsymbol{M};K}$ are monic, it is easily seen that $(\zeta^{-1}_{\boldsymbol{M}})^\bullet_{;K}$ and $(\zeta^{-1}_{\boldsymbol{M}})^{\mathrm{int}}_{;O}$ are inverses to $\zeta^\bullet_{\boldsymbol{M};K}$ and $\zeta^{\mathrm{int}}_{\boldsymbol{M};O}$, which are therefore isomorphisms.

Finally, in the case that $\zeta_{\boldsymbol{M}}$ is a component of a natural transformation $\zeta : \mathscr{B} \dot\to \mathscr{A}$, Eq. (6.1) holds by Prop. 3.8. □

Lemma 6.7 *Suppose $\zeta : \mathscr{B} \dot\to \mathscr{A}$ and that there exist subobjects $\psi_i : \boldsymbol{M}_i \to \boldsymbol{M}$ $(i \in I)$ such that $\bigvee_{i\in I} \mathscr{A}(\psi_i)$ and all the $\zeta_{\boldsymbol{M}_i}$ are isomorphisms. Then $\zeta_{\boldsymbol{M}}$ and $\bigvee_{i\in I} \mathscr{B}(\psi_i)$ are isomorphisms.*

Proof: Consider, for each $i \in I$, the diagram

$$\begin{tikzcd}
\mathscr{A}(\boldsymbol{M}_i) \arrow[r] \arrow[d, "\zeta^{-1}_{\boldsymbol{M}_i}"'] & \displaystyle\bigvee_{i\in I} \mathscr{A}(\boldsymbol{M}_i) \arrow[rr, "\bigvee_{i\in I} \mathscr{A}(\psi_i)"] \arrow[d, dotted, "\xi"'] & & \mathscr{A}(\boldsymbol{M}) \arrow[d, "\mathrm{id}_{\mathscr{A}(\boldsymbol{M})}"'] \\
\mathscr{B}(\boldsymbol{M}_i) \arrow[r] & \displaystyle\bigvee_{i\in I} \mathscr{B}(\boldsymbol{M}_i) \arrow[r, "\bigvee_{i\in I} \mathscr{B}(\psi_i)"] & \mathscr{B}(\boldsymbol{M}) \arrow[r, "\zeta_{\boldsymbol{M}}"] & \mathscr{A}(\boldsymbol{M})
\end{tikzcd}$$

in which the unlabelled morphisms are the canonical inclusions associated with the join. Thus the two horizontal morphisms on the top line compose to give $\mathscr{A}(\psi_i)$, and the left two horizontal morphisms on the bottom line compose to give $\mathscr{B}(\psi_i)$. The outer portion of the diagram therefore commutes because ζ is natural and the universal property of the union

induces a unique morphism ξ such that every such diagram commutes in full. Considering the right-hand rectangle, it is evident that $\zeta_{\boldsymbol{M}}$ and $\bigvee_{i\in I}\mathscr{B}(\psi_i)$ have inverses

$$\zeta_{\boldsymbol{M}}^{-1} = \left(\bigvee_{i\in I}\mathscr{B}(\psi_i)\right)\circ\xi\circ\left(\bigvee_{i\in I}\mathscr{A}(\psi_i)\right)^{-1}$$

$$\left(\bigvee_{i\in I}\mathscr{B}(\psi_i)\right)^{-1} = \xi\circ\left(\bigvee_{i\in I}\mathscr{A}(\psi_i)\right)^{-1}\circ\zeta_{\boldsymbol{M}};$$

hence they are isomorphisms. $\square$

Both the previous results hold regardless of whether $\mathscr{A}$ and $\mathscr{B}$ are dynamically local (indeed, Lemma 6.7 does not even use the timeslice property). Given the additional assumption we can use Lem. 6.6 to prove:

Proposition 6.8 *Suppose $\mathscr{A}$ and $\mathscr{B}$ are dynamically local and $\zeta:\mathscr{B}\dot\to\mathscr{A}$. Suppose in addition that $\zeta_{\boldsymbol{N}}$ is an isomorphism for some $\boldsymbol{N}$. Then $\zeta_{\boldsymbol{M}}$ is an isomorphism for all $\boldsymbol{M}$ for which there is a morphism $\boldsymbol{M}\to\boldsymbol{N}$.*

Proof: We have a diagram

$$\begin{tikzcd}
\mathscr{B}(\boldsymbol{M}) \arrow[rrr, "\zeta_{\boldsymbol{M}}"] \arrow[dr, "\mathscr{B}(\psi)"] \arrow[dd, "\cong"'] & & & \mathscr{A}(\boldsymbol{M}) \arrow[dl, "\mathscr{A}(\psi)"'] \arrow[dd, "\cong"] \\
& \mathscr{B}(\boldsymbol{N}) \arrow[r, "\zeta_{\boldsymbol{N}}"] & \mathscr{A}(\boldsymbol{N}) & \\
\mathscr{B}^{\mathrm{dyn}}(\psi(\boldsymbol{M});\boldsymbol{N}) \arrow[ur, "\beta^{\mathrm{dyn}}_{\boldsymbol{N};\psi(\mathcal{M})}"] \arrow[rrr, "\zeta_{\boldsymbol{N};\psi(\boldsymbol{M})}"] & & & \mathscr{A}^{\mathrm{dyn}}(\psi(\boldsymbol{M});\boldsymbol{N}) \arrow[ul, "\alpha^{\mathrm{dyn}}_{\boldsymbol{N};\psi(\mathcal{M})}"']
\end{tikzcd}$$

in which the two vertical isomorphisms arise because $\mathscr{B}(\psi)\cong\beta^{\mathrm{dyn}}_{\boldsymbol{N};\psi(\mathcal{M})}$ and $\mathscr{A}(\psi)\cong\alpha^{\mathrm{dyn}}_{\boldsymbol{N};\psi(\mathcal{M})}$ by dynamical locality, whereupon the side triangles commute. The upper trapezium commutes by naturality of ζ and the lower trapezium by Lemma 6.6, which also entails that $\zeta_{\boldsymbol{N};\psi(\boldsymbol{M})}$ is an isomorphism. Thus the diagram commutes in full, implying that $\zeta_{\boldsymbol{M}}$ is an isomorphism by commutativity of the outer rectangle. $\square$

In addition, we will use the following simple result (here dynamical locality is not assumed):

Proposition 6.9 *Suppose $\zeta:\mathscr{B}\dot\to\mathscr{A}$. If $\psi:\boldsymbol{M}\to\boldsymbol{N}$ is a Cauchy morphism, then $\zeta_{\boldsymbol{M}}$ is an isomorphism if and only if $\zeta_{\boldsymbol{N}}$ is an isomorphism.*

Proof: We have $\mathscr{A}(\psi)\circ\zeta_{\boldsymbol{M}} = \zeta_{\boldsymbol{N}}\circ\mathscr{B}(\psi)$, with $\mathscr{A}(\psi)$ and $\mathscr{B}(\psi)$ isomorphisms. If $\zeta_{\boldsymbol{N}}$ is an isomorphism then $\mathscr{B}(\psi)^{-1}\circ\zeta_{\boldsymbol{N}}^{-1}\circ\mathscr{A}(\psi)$ is inverse for $\zeta_{\boldsymbol{M}}$, and hence $\zeta_{\boldsymbol{M}}$ is an isomorphism. Similarly $\mathscr{B}(\psi)\circ\zeta_{\boldsymbol{M}}^{-1}\circ\mathscr{A}(\psi)^{-1}$ is inverse to $\zeta_{\boldsymbol{N}}$ if $\zeta_{\boldsymbol{M}}$ is an isomorphism. □

Given the above preparation, we may now state and prove our main result of this section: namely that the dynamically local theories have the SPASs property.

Theorem 6.10 *Suppose $\mathscr{A}$ and $\mathscr{B}$ dynamically local theories and $\zeta : \mathscr{B} \stackrel{.}{\to} \mathscr{A}$. If $\zeta_{\boldsymbol{M}}$ is an isomorphism for some spacetime $\boldsymbol{M}$ then ζ is a natural isomorphism.*

Proof: Given that $\zeta_{\boldsymbol{M}}$ is an isomorphism, Prop. 6.8 entails that $\zeta_{\boldsymbol{D}}$ is an isomorphism for any multi-diamond spacetime $\boldsymbol{D} \to \boldsymbol{M}$. Now let $\boldsymbol{D}'$ be any other multi-diamond spacetime with the same number of components as $\boldsymbol{D}$; as $\boldsymbol{D}$ and $\boldsymbol{D}'$ have oriented-diffeomorphic Cauchy surfaces, they are linked by a chain of Cauchy morphisms by as shown in Prop. 2.4. Using Props. 6.9 and 6.8, we may conclude that $\zeta_{\boldsymbol{D}'}$ is also an isomorphism. As $\boldsymbol{M}$ contains multi-diamonds with any finite number of components, it follows that $\zeta_{\boldsymbol{D}'}$ is an isomorphism for every multi-diamond spacetime $\boldsymbol{D}'$.

Now let $\boldsymbol{M}'$ be an arbitrary spacetime; as $\mathscr{A}$ is dynamically local, we may deduce that $\zeta_{\boldsymbol{M}'}$ is an isomorphism using Theorem 6.3(b) and the remark thereafter, in conjunction with Lemma 6.7. □

Thus for any dynamically local theory $\mathscr{A}$, there is no simpler dynamically local theory that could account for the physics in any particular spacetime. In this sense, dynamical locality therefore ensures that $\mathscr{A}$ has the same physical content in all spacetimes. Examples of the type presented in Section 4 include cases where $\mathscr{A}$ (resp., $\mathscr{B}$) is dynamically local, but $\mathscr{B}$ (resp., $\mathscr{A}$) is not and where there is a partial isomorphism $\mathscr{B} \to \mathscr{A}$ that is not an isomorphism. Let us note that much of the argument depends largely on the additivity property (that is a consequence of dynamical locality). The exception is Prop. 6.8, where additivity seems to be insufficient, and one requires the stronger dynamical locality assumption.

To conclude this section, we consider the consequences of dynamical locality for ordinary diagonal theories.

Theorem 6.11 *Suppose φ_Δ is an ordinary diagonal theory such that φ_Δ and every $\varphi(\boldsymbol{M})$ are dynamically local. Then (a) for every morphism $\psi : \boldsymbol{M} \to \boldsymbol{N}$, $\varphi(\psi)$ is a natural isomorphism; (b) φ_Δ is gauge-equivalent to any $\varphi(\boldsymbol{M})$. In particular, if $\mathrm{Aut}(\varphi(\boldsymbol{M}))$ is trivial, then φ_Δ is equivalent to each $\varphi(\boldsymbol{M})$.*

Proof: (a) We have

$$\varphi_\Delta(\psi) \cong \alpha^{(\varphi_\Delta)\mathrm{kin}}_{\boldsymbol{N};\psi(\boldsymbol{M})} \cong \alpha^{(\varphi_\Delta)\mathrm{dyn}}_{\boldsymbol{N};\psi(\boldsymbol{M})} \cong \alpha^{(\varphi(\boldsymbol{N}))\mathrm{dyn}}_{\boldsymbol{N};\psi(\boldsymbol{M})} \cong \alpha^{(\varphi(\boldsymbol{N}))\mathrm{kin}}_{\boldsymbol{N};\psi(\boldsymbol{M})} \cong \varphi(\boldsymbol{N})(\psi)$$

using dynamical locality of φ_Δ and $\varphi(\boldsymbol{N})$ and Thm. 5.7 (expressed in subobject language). Hence $\varphi(\boldsymbol{N})(\psi)\circ\varphi(\psi)_{\boldsymbol{M}} \cong \varphi(\boldsymbol{N})(\psi)$ and as $\varphi(\boldsymbol{N})(\psi)$ is monic, $\varphi(\psi)_{\boldsymbol{M}}$ is an isomorphism. As both $\varphi(\boldsymbol{M})$ and $\varphi(\boldsymbol{N})$ are dynamically local, Theorem 6.10 entails that $\varphi(\psi)$ is a natural isomorphism.

(b) Writing $\boldsymbol{M}_0$ for Minkowski space, for each $\boldsymbol{M}$ we may choose a chain of morphisms as in Prop. 2.6

$$\boldsymbol{M}_0 \leftarrow \boldsymbol{M}_1 \rightarrow \boldsymbol{M}_2 \leftarrow \boldsymbol{M}_3 \rightarrow \boldsymbol{M}$$

and use part (a) four times, composing the corresponding natural isomorphisms or their inverses, to obtain a natural isomorphism $\zeta_{\boldsymbol{M}} : \varphi(\boldsymbol{M}_0) \stackrel{.}{\to} \varphi(\boldsymbol{M})$. Then for each $\psi : \boldsymbol{M} \to \boldsymbol{N}$, $\eta(\psi) := \zeta_{\boldsymbol{N}}^{-1} \circ \varphi(\psi) \circ \zeta_{\boldsymbol{M}}$ is an automorphism of $\varphi(\boldsymbol{M}_0)$. It is obvious that $\eta(\psi \circ \psi') = \eta(\psi) \circ \eta(\psi')$ and $\eta(\mathrm{id}_{\boldsymbol{M}}) = \mathrm{id}_{\varphi(\boldsymbol{M}_0)}$. Thus $\eta \in \mathrm{Funct}(\mathsf{Loc}_0, \mathrm{Aut}(\varphi(\boldsymbol{M}_0)))$ [with the automorphism group regarded as a category] and we have

$$\begin{aligned}\varphi_\Delta(\psi) \circ (\zeta_{\boldsymbol{M}})_{\boldsymbol{M}} &= \varphi(\boldsymbol{N})(\psi) \circ \varphi(\psi) \circ (\zeta_{\boldsymbol{M}})_{\boldsymbol{M}} = \varphi(\boldsymbol{N})(\psi) \circ (\zeta_{\boldsymbol{M}})_{\boldsymbol{N}} \circ \eta(\psi)_{\boldsymbol{M}} \\ &= (\zeta_{\boldsymbol{N}})_{\boldsymbol{N}} \circ \varphi(\boldsymbol{M}_0)(\psi) \circ \eta(\psi)_{\boldsymbol{M}}.\end{aligned}$$

Thus the morphisms $(\zeta_{\boldsymbol{M}})_{\boldsymbol{M}}$ form the components of a natural transformation up to the twisting η.

Finally, if the automorphism group is trivial, $\eta(\psi)$ is an identity for all ψ and the $(\zeta_{\boldsymbol{M}})_{\boldsymbol{M}}$ become components of a natural isomorphism $\hat{\zeta} : \varphi(\boldsymbol{M}_0) \stackrel{.}{\to} \varphi_\Delta$. □

This result raises the interesting issue of how much freedom is available through choice of η, which can be regarded as a cohomological issue. If $\mathrm{Aut}(\varphi(\boldsymbol{M}_0))$ is nontrivial, we can see that inequivalent diagonal theories can be constructed in the following way. Label every homeomorphism equivalence class $[\Sigma]$ of compact connected Riemannian manifold by an element $g_{[\Sigma]}$ of $\mathrm{Aut}(\varphi(\boldsymbol{M}_0))$, and for each morphism $\psi : \boldsymbol{M} \to \boldsymbol{N}$ in Loc_0 define $\eta(\psi)$ to be trivial except in the case that $\boldsymbol{M}$ has noncompact Cauchy surfaces and $\boldsymbol{N}$ has compact Cauchy surface, in which case we set $\eta(\psi) = g_{[\Sigma(\boldsymbol{N})]}$. It is clear that this defines a functor into $\mathrm{Aut}(\varphi(\boldsymbol{M}_0))$.

6.3 A no-go theorem for natural states

To illustrate the significance of the dynamical locality assumption, we prove a model-independent no-go theorem for assignments of a natural choice of preferred state of a QFT in all spacetimes. This brings to sharper form an argument sketched in BFV and [32] for the free scalar field; essentially it shows that a preferred state is essentially incompatible with quantum field theory.

Unlike the results above, this result is specific to situations in which Phys is a category of $*$-algebras [including C^*-Alg]. We realise the subobjects $\mathscr{A}^{\bullet/\mathrm{dyn}/\mathrm{kin}}(\boldsymbol{M};O)$ as subalgebras of $\mathscr{A}(\boldsymbol{M})$ throughout. In this context, a state of the theory in spacetime $\boldsymbol{M}$ is a normalised positive linear functional on the algebra $\mathscr{A}(\boldsymbol{M})$; the space of all states is denoted $\mathscr{A}(\boldsymbol{M})^*_{+,1}$. The result is stated in LCT but has an obvious analogue in LCT_0.

Definition 6.12 *A* natural state *of a theory $\mathscr{A}$ in LCT is an assignment $\mathsf{Loc} \ni \boldsymbol{M} \mapsto \omega_{\boldsymbol{M}} \in \mathscr{A}(\boldsymbol{M})^*_{+,1}$ such that $\mathscr{A}(\psi)^*\omega_{\boldsymbol{N}} = \omega_{\boldsymbol{M}}$ for all morphisms $\psi : \boldsymbol{M} \to \boldsymbol{N}$.*

Theorem 6.13 *Suppose $\mathscr{A}$ is a dynamically local theory in LCT, and has a natural state $(\omega_{\boldsymbol{M}})_{\boldsymbol{M}\in\mathsf{Loc}}$. If there is a spacetime $\boldsymbol{M}$ with noncompact Cauchy surfaces such that $\omega_{\boldsymbol{M}}$*

induces a faithful GNS representation with the Reeh–Schlieder property [i.e., the GNS vector corresponding to $\omega_{\boldsymbol{M}}$ is cyclic for the induced representation of $\mathscr{A}(\boldsymbol{M}|_O)$ for all relatively compact $O \in \mathscr{O}_0(\boldsymbol{M})$], then the relative Cauchy evolution is trivial in $\boldsymbol{M}$. If, additionally, $\mathscr{A}$ obeys extended locality, then $\mathscr{A}$ is equivalent to the trivial theory $\mathscr{I}$.

Proof: Let $\boldsymbol{M}$ be as in the statement of the theorem. As the relative Cauchy evolution is a composition of (inverses of) morphisms $\mathscr{A}(\psi)$, we have $\omega_{\boldsymbol{M}} \circ \mathrm{rce}_{\boldsymbol{M}}[\boldsymbol{h}] = \omega_{\boldsymbol{M}}$ for each $\boldsymbol{M}$ and all $\boldsymbol{h} \in H(\boldsymbol{M})$. Consequently, in the GNS representation $\pi_{\boldsymbol{M}}$ induced by $\omega_{\boldsymbol{M}}$, the relative Cauchy evolution may be unitarily implemented as

$$\pi_{\boldsymbol{M}}(\mathrm{rce}_{\boldsymbol{M}}[\boldsymbol{h}]A) = U_{\boldsymbol{M}}[\boldsymbol{h}]\pi_{\boldsymbol{M}}(A)U_{\boldsymbol{M}}[\boldsymbol{h}]^{-1}$$

for unitaries $U_{\boldsymbol{M}}[\boldsymbol{h}]$ defined by $U_{\boldsymbol{M}}[\boldsymbol{h}]\pi_{\boldsymbol{M}}(A)\Omega_{\boldsymbol{M}} = \pi_{\boldsymbol{M}}(\mathrm{rce}_{\boldsymbol{M}}[\boldsymbol{h}]A)\Omega_{\boldsymbol{M}}$, leaving the GNS vector $\Omega_{\boldsymbol{M}}$ invariant.

Now let $\boldsymbol{h} \in H(\boldsymbol{M})$ and choose a nonempty relatively compact $O \in \mathscr{O}_0(\boldsymbol{M})$ such that $O \subset (\mathrm{supp}\,\boldsymbol{h})^\perp$ (here we use the noncompactness of the Cauchy surfaces). Then by Prop. 3.5 and Lem. 3.1, we have

$$\mathrm{rce}_{\boldsymbol{M}}[\boldsymbol{h}] \circ \alpha^{\mathrm{kin}}_{\boldsymbol{M};O} = \alpha^{\mathrm{kin}}_{\boldsymbol{M};O}$$

and hence that

$$U_{\boldsymbol{M}}[\boldsymbol{h}]\pi_{\boldsymbol{M}}(\mathscr{A}(\iota_{\boldsymbol{M};O})A)\Omega_{\boldsymbol{M}} = \pi_{\boldsymbol{M}}(\mathscr{A}(\iota_{\boldsymbol{M};O})A)\Omega_{\boldsymbol{M}}$$

for all $A \in \mathscr{A}(\boldsymbol{M}|_O)$. Using the Reeh–Schlieder assumption on $\omega_{\boldsymbol{M}}$ we may deduce that $U_{\boldsymbol{M}}[\boldsymbol{h}]$ agrees with the identity operator on a dense set and hence $U_{\boldsymbol{M}}[\boldsymbol{h}] = \mathbf{1}_{\mathscr{H}_{\boldsymbol{M}}}$ for all $\boldsymbol{h} \in H(\boldsymbol{M})$.

As the representation $\pi_{\boldsymbol{M}}$ is assumed faithful, the relative Cauchy evolution is trivial on $\mathscr{A}(\boldsymbol{M})$ as claimed. Consequently, $\mathscr{A}^\bullet(\boldsymbol{M};K) = \mathscr{A}(\boldsymbol{M})$ for all compact sets K and hence by dynamical locality $\mathscr{A}^{\mathrm{kin}}(\boldsymbol{M};O) = \mathscr{A}^{\mathrm{dyn}}(\boldsymbol{M};O) = \mathscr{A}(\boldsymbol{M})$ for each nonempty $O \in \mathscr{O}(\boldsymbol{M})$.

Now consider two causally disjoint nonempty $O_1, O_2 \in \mathscr{O}(\boldsymbol{M})$ (it suffices that they are each connected). It is clear that $\mathscr{A}$ can obey extended locality only if $\mathscr{A}(\boldsymbol{M}) = \mathbb{C}\mathbf{1}_{\mathscr{A}(\boldsymbol{M})}$. (The same would also be true if the $\mathscr{A}^{\mathrm{kin}}(\boldsymbol{M};O_i)$ are required to be algebraically independent: otherwise we can find a linearly independent set $\{\mathbf{1}_{\mathscr{A}(\boldsymbol{M})}, A\}$ common to the two algebras, whose list of products are of course linearly dependent.)

Thus the subtheory embedding $\mathscr{I}_{\mathscr{A}} : \mathscr{I} \dot{\to} \mathscr{A}$ is an isomorphism in spacetime $\boldsymbol{M}$. As both $\mathscr{I}$ and $\mathscr{A}$ are assumed dynamically local, it follows from Theorem 6.10 that $\mathscr{I}_{\mathscr{A}}$ is a natural isomorphism. □

We remark that the assumption of commutation at spacelike separation (in place of extended locality) results in $\mathscr{A}(\boldsymbol{M})$ being abelian, from which we can deduce that $\mathscr{A}(\boldsymbol{N})$ is abelian if $\boldsymbol{N}$ is any truncated multi-diamond spacetime, or any spacetime in which the truncated multi-diamonds form a directed net.

7 Conclusion

We conclude with a brief discussion of further work and related approaches. First, now that the basic framework has been established, it is necessary to show that familiar models satisfy dynamical locality. As already mentioned, we show in [28] that the minimally coupled free scalar field is dynamically local for nonzero mass, and that the failure of dynamical locality at zero mass is understood as an expression of the gauge symmetry. Once this is taken into account the massless theory is again dynamically local, with the single exception of the two-dimensional theory on Loc. Work is under way on other models, including the algebra of Wick products.

Second, we again emphasise that we do not expect that the two principles S1 and S2 described in the introduction completely characterise what a notion of SPASs should be. For example, it is conceivable that there are (as yet unknown) dynamically local theories that one might not wish to regard as representing the same physics in all spacetimes; in that case, it would be clear that S1 and S2 are insufficient and that further conditions should be imposed. Furthermore our discussion is conducted for the most part at the level of local observables. Even in the algebraic approach to quantum field theory in curved spacetimes there are several levels of description and the present work addresses only those aspects that are independent of choices of state spaces which can bring in properties deriving from the global structure of spacetime. In addition, it is possible that the formulation of dynamical locality can be refined further. For example, one might base the theory on the requirement that $\alpha^{\bullet}_{\boldsymbol{M};K}$ should be isomorphic to any intersection $\bigwedge_n \alpha^{\rm kin}_{\boldsymbol{M};O_n}$ where O_n is an outer approximating sequence to K.

Finally, we conclude with some remarks that may help to clarify the relation of the present work to other approaches studying the interplay of covariance, locality and dynamics in abstract (operator-algebraic) quantum field theory. If our setting is specialized to the case that Phys is C^*-Alg, the category of unital C^* -algebras, then our discussion remains purely at the C^*-algebraic level, in that we do not discuss special classes of states or their GNS representations, from which, in a next step, the C^*-algebraic setting would be taken to the von Neumann-algebraic level. This step, together with the analysis of distinguished states and their induced representations, is one of the central issues in the model-independent approach to quantum field theory, as is laid out in [30], and other work devoted to the relations between covariance, locality and dynamics is mostly tied to distinguished states, often the vacuum state in Minkowski spacetime. Some authors have attempted to derive a concept of dynamical localization of observables for quantum field theory in Minkowski spacetime, making use of the properties of the vacuum representation [39, 37]; however this concept of dynamical localization is different from ours. Another major theme in operator-algebraic quantum field theory is the concept of "geometric modular action" [13] which has at its roots the famous Bisognano–Wichmann theorem (see [30], and references cited there). This theorem says that the Tomita–Takesaki modular objects corresponding to von Neumann algebras of observables localized in special regions, and to the vacuum vector, carry geometrical significance. In fact, in some situations one can gain the full local net structure and covariance group from such modular objects [13, 53]. This

is of interest as the modular objects also encode dynamical information [30], and in some works, this dynamical information has been related to concepts of locality and covariance [17, 10, 15, 14]. While these cited works are not directly related to the approach taken in the present article, they also focus on the relation between covariance, locality and dynamics. Closer connections between the cited works and the present article may possibly be revealed once our setup can suitably be extended at the von Neumann algebraic level, incorporating distinguished classes of states.

CJF thanks the Research Academy Leipzig for financial support and the Institute for Theoretical Physics, University of Leipzig, for hospitality at various stages of this work. In addition, CJF thanks the Insitute for Theoretical Physics, University of Göttingen for hospitality and the participants and organisers of the meetings AQFT50 (Göttingen, July–August 2009) and Quantum Field Theory and Gravity (Regensburg, September–October 2010) at which preliminary versions of this work were presented. CJF also thanks Henning Bostelmann, Matthew Ferguson, Klaus Fredenhagen and Ko Sanders for useful conversations at various stages during this work. RV thanks the Department of Mathematics, University of York, for hospitality and financial support.

A Geometrical lemmas

A.1 Cauchy morphisms

Proposition A.1 *Suppose $\boldsymbol{M} \in \mathsf{Loc}_0$ admits a compact Cauchy surface Σ. If $\boldsymbol{M} \xrightarrow{\psi} \boldsymbol{N}$ in Loc_0 then ψ is Cauchy.*

Proof. Using homeomorphism equivalence of Cauchy surfaces in $\boldsymbol{M}$ [44, Cor. 14.32] and [3, Thm. 1.1], we may assume without loss of generality that Σ is a smooth spacelike Cauchy surface, which is connected [44, Prop. 14.31], compact and embedded in $\boldsymbol{M}$. As ψ is an isometric embedding, $\psi(\Sigma)$ is (in particular) a smoothly immersed spacelike submanifold of $\boldsymbol{N}$ that is also compact and connected as a result of the properties of Σ. Theorem 1 of [16] then entails that $\psi(\Sigma)$ is an acausal Cauchy surface[18] of $\boldsymbol{N}$, so ψ is Cauchy. □

Lemma A.2 *Suppose $\psi : \boldsymbol{M} \to \boldsymbol{N}$ is Cauchy. Then if Σ is any Cauchy surface of $\boldsymbol{M}$, $\psi(\Sigma)$ is a Cauchy surface of $\boldsymbol{N}$.*

Proof: Any inextendible timelike curve $\gamma : \mathbb{R} \to \boldsymbol{N}$ in $\boldsymbol{N}$ enters $\psi(\boldsymbol{M})$, and $I = \gamma^{-1}(\psi(\boldsymbol{M}))$ is open and connected by causal convexity of the embedding. We therefore obtain a timelike curve $\hat{\gamma} : I \to \boldsymbol{M}$ so that $\psi \circ \hat{\gamma} = \gamma|_I$. Now $\hat{\gamma}$ has no endpoint in $\boldsymbol{M}$ and is therefore inextendible; accordingly it intersects Σ exactly once. Hence $\gamma|_I$ intersects $\psi(\Sigma)$ exactly once and so the same is true of γ. □

Lemma A.3 *The composite of Cauchy morphisms is Cauchy.*

[18] Note that 'acausal' is included in the definition of Cauchy surface in [16].

Proof: If $\varphi : \boldsymbol{L} \to \boldsymbol{M}$ and $\psi : \boldsymbol{M} \to \boldsymbol{N}$ are Cauchy and Σ is a Cauchy surface of $\boldsymbol{L}$, then we apply Lem. A.2 successively, to show that $\varphi(\Sigma)$ is a Cauchy surface of $\boldsymbol{M}$ and hence $(\psi \circ \varphi)(\Sigma)$ is a Cauchy surface of $\boldsymbol{N}$. Hence $\psi \circ \varphi$ is Cauchy. $\square$

We now give two proofs deferred from section 2.

Proof of Prop. 2.2: We are given a Cauchy morphism $\psi : \boldsymbol{M} \to \boldsymbol{N}$ in Loc_0 or Loc, and must prove that $\psi(\boldsymbol{M})$ contains a smooth, spacelike and acausal Cauchy surface for $\boldsymbol{N}$ and that the smooth spacelike Cauchy surfaces of $\boldsymbol{M}$ and $\boldsymbol{N}$ are oriented-diffeomorphic.

By virtue of [3, Thm 1.1] $\boldsymbol{M}$ has a smooth spacelike Cauchy surface Σ; Lem. A.2 shows that the smooth spacelike surface $\Sigma' = \psi(\Sigma)$ is a Cauchy surface for $\boldsymbol{N}$, and is therefore acausal by [44, Lem. 14.42]. Putting $\boldsymbol{M}$ and $\boldsymbol{N}$ into normal form, we may construct oriented-diffeomorphisms $\rho : \mathbb{R} \times \Sigma \to \boldsymbol{M}$, $\rho' : \mathbb{R} \times \Sigma' \to \boldsymbol{N}$ (with the canonical orientations and other properties discussed in section 2), thus giving a smooth map $\Psi = \mathrm{pr}_{\Sigma'} \circ (\rho')^{-1} \circ \psi \circ \rho_0 : \Sigma \to \Sigma'$, where $\mathrm{pr}_{\Sigma'}$ is the projection onto Σ' and $\rho_0(\cdot) = \rho(0, \cdot)$.

Now Ψ is an immersion (and hence also a submersion) because the kernel of $(\mathrm{pr}_{\Sigma'} \circ (\rho')^{-1})_*$ is timelike while the image of $(\psi \circ \rho_0)_*$ is spacelike; it is also injective ($\mathrm{pr}_{\Sigma'} \circ (\rho')^{-1}$ identifies points only if they are connected by a timelike curve, while $\rho_0(\Sigma)$ is achronal and $\psi(\boldsymbol{M})$ is causally convex) and surjective (by definition of $\Sigma' = \psi(\Sigma)$ and because $\mathrm{pr}_{\Sigma'} \circ (\rho')^{-1} \circ \rho'_0 = \mathrm{id}_{\Sigma'}$). Accordingly, Ψ is a diffeomorphism (see, e.g., [40, Thm 7.15]) that preserves orientations because ψ preserves orientation and time-orientation. In particular, Ψ is a homeomorphism and so all Cauchy surfaces of $\boldsymbol{M}$ and $\boldsymbol{N}$ are homeomorphic. $\square$

Proof of 2.4 (converse): Our argument is a slight elaboration and variant of that in [29] in order to incorporate detail on orientations. We also take the opportunity to simplify the argument slightly, while also being more specific on some details. We suppose $\boldsymbol{M}$ and $\boldsymbol{N}$ have oriented-diffeomorphic smooth spacelike Cauchy surfaces Σ and Σ' with canonical orientations $\mathfrak{w}$ and $\mathfrak{w}'$. Using any oriented-diffeomorphism between Σ and Σ' we may put both $\boldsymbol{M}$ and $\boldsymbol{N}$ into normal form on $\mathbb{R} \times \Sigma$ equipped with the orientation $dt \wedge \mathfrak{w}$ by means of oriented-diffeomorphisms $\rho_{\boldsymbol{M}} : \mathbb{R} \times \Sigma \to \boldsymbol{M}$ and $\rho_{\boldsymbol{N}} : \mathbb{R} \times \Sigma \to \boldsymbol{N}$. The two pulled back metrics on $\mathbb{R} \times \Sigma$ may be written as

$$\rho_{\boldsymbol{M}}^* \boldsymbol{g}_{\boldsymbol{M}} = \beta_{\boldsymbol{M}} dt \otimes dt - \boldsymbol{h}_t, \qquad \rho_{\boldsymbol{N}}^* \boldsymbol{g}_{\boldsymbol{N}} = \beta_{\boldsymbol{N}} dt \otimes dt - \boldsymbol{k}_t$$

where $\beta_{\boldsymbol{M}}, \beta_{\boldsymbol{N}} \in C^\infty(\mathbb{R} \times \Sigma)$ are strictly positive and $\boldsymbol{h}_t$ and $\boldsymbol{k}_t$ are smooth Riemannian metrics on Σ depending smoothly on t. One may find smooth positive functions $K, H \in C^\infty(\mathbb{R} \times \Sigma)$ such that $\boldsymbol{k}_{t,\sigma} \geq K(t,\sigma) \boldsymbol{h}_{t,\sigma}$ and $\boldsymbol{h}_{t,\sigma} \geq H(t,\sigma) \boldsymbol{k}_{t,\sigma}$ as quadratic forms.[19] Fixing $t_0 > 0$, let $F = (t_0, \infty) \times \Sigma$ and $P = (-\infty, -t_0) \times \Sigma$ and choose any nonnegative $\chi \in C^\infty(\mathbb{R})$ such that χ equals unity on F and vanishes on P. Construct a metric

$$\boldsymbol{g} = \beta dt \otimes dt - (\chi \boldsymbol{h}_t + (1 - \chi) \boldsymbol{k}_t)$$

where β is chosen to be a smooth positive function such that

$$\beta \leq (\chi + (1 - \chi) K) \beta_{\boldsymbol{M}}$$

[19] E.g., use $K = [(\boldsymbol{h}_t)^i{}_j (\boldsymbol{h}_t)^j{}_i]^{-1/2}$ with $\boldsymbol{k}_t$ used to raise indices, and the analogous expression for H.

on $t > -\frac{1}{2}t_0$, with equality for $t \geq t_0$, and

$$\beta \leq (1 - \chi + \chi H)\beta_{\boldsymbol{N}}$$

on $t < \frac{1}{2}t_0$, with equality for $t \leq -t_0$. Then it is easily seen that every $\boldsymbol{g}$-causal curve is $\rho_{\boldsymbol{M}}^* \boldsymbol{g}_{\boldsymbol{M}}$-causal in $(-\frac{1}{2}t_0, \infty) \times \Sigma$ and $\rho_{\boldsymbol{N}}^* \boldsymbol{g}_{\boldsymbol{N}}$-causal in $(-\infty, \frac{1}{2}t_0) \times \Sigma$. But these metrics are globally hyperbolic, so every inextendible $\boldsymbol{g}$-timelike curve intersects each $\{t\} \times \Sigma$ surface exactly once. Accordingly, $\mathbb{R} \times \Sigma$, with the metric $\boldsymbol{g}$, orientation $dt \wedge \mathfrak{w}$ and time-orientation so that $\partial/\partial t$ is future-pointing, is a globally hyperbolic spacetime in Loc (or Loc_0 as appropriate), which we denote $\boldsymbol{I}$. The metric $\boldsymbol{g}$ clearly coincides with $\rho_{\boldsymbol{M}}^* \boldsymbol{g}_{\boldsymbol{M}}$ in F and with $\rho_{\boldsymbol{N}}^* \boldsymbol{g}_{\boldsymbol{N}}$ on P.

Finally, the regions F and P are open globally hyperbolic subsets of $\boldsymbol{I}$ containing Cauchy surfaces of $\boldsymbol{I}$ and their images $\rho_{\boldsymbol{M}}(F)$ and $\rho_{\boldsymbol{N}}(P)$ evidently contain Cauchy surfaces for $\boldsymbol{M}$ and $\boldsymbol{N}$. Setting $\boldsymbol{F} = \boldsymbol{I}|_F$ and $\boldsymbol{P} = \boldsymbol{I}|_P$, we then have a diagram of the form (2.1)

$$\boldsymbol{M} \leftarrow \boldsymbol{F} \rightarrow \boldsymbol{I} \leftarrow \boldsymbol{P} \rightarrow \boldsymbol{N},$$

with the canonical inclusions $\iota_{\boldsymbol{I};\boldsymbol{F}}$ and $\iota_{\boldsymbol{I};\boldsymbol{P}}$ providing the inner Cauchy morphisms and the restrictions $\rho_{\boldsymbol{M}}|_F$ and $\rho_{\boldsymbol{N}}|_P$ as the outer two Cauchy morphisms. $\square$.

A.2 Covariance of hyperbolic perturbations

Next, we turn to a number of results used in the discussion of relative Cauchy evolution in Sect. 3.4. We recall that the chronological future(+)/past(−) $I_{\boldsymbol{M}}^{\pm}(p)$ of p consists of all points (excluding p) that can be reached from p along a future/past-directed piecewise smooth timelike curve in $\boldsymbol{M}$; by smoothing results such as [45, Prop. 2.23] we obtain the same set if we only admit smooth timelike curves (which may even be chosen to be geodesic near their endpoints). Similarly, the causal future/past $J_{\boldsymbol{M}}^{\pm}(p)$ consists of all points (including p) that can be reached from p by future/past directed piecewise smooth (or, equivalently, smooth) causal curves. Note that any causal curve is confined to a single connected component of the spacetime. For a subset $S \subset \boldsymbol{M}$ we define $J_{\boldsymbol{M}}^{\pm}(S) = \bigcup_{p \in S} J_{\boldsymbol{M}}^{\pm}(p)$ etc. Extensive use will be made of the fact that globally hyperbolic spacetimes are causally simple: for every compact set K, the sets $J_{\boldsymbol{M}}^{\pm}(K)$ are closed (see, e.g., Prop. 6.6.1 in [31]; Theorem 8.3.11 in [52]).

For any subset $S \subset \boldsymbol{M}$ we define the future(+)/past(−) Cauchy development $D^{\pm}(S)$ of S to be the set of points p such that every past/future-inextendible piecewise smooth causal curve through p intersects S; $D_{\boldsymbol{M}}(S) = D_{\boldsymbol{M}}^{+}(S) \cup D_{\boldsymbol{M}}^{-}(S)$. If S is either achronal or closed, we may replace 'piecewise smooth' by 'smooth' without loss, but more generally, this can result in a different set.

Proof of Lemma 3.2: If $\boldsymbol{M}$ is connected, this is immediate from the special case $K = \operatorname{supp} \boldsymbol{h}$ of the following result, Lemma A.4. If $\boldsymbol{M}$ has more than one connected component, the result follows by applying Lemma A.4 to each component. $\square$

Lemma A.4 *Let K be a compact subset of the underlying manifold $\mathcal{M}$ of $\boldsymbol{M} \in \mathsf{Loc}_0$ and define $\mathcal{M}^{\pm} = \mathcal{M} \setminus J_{\boldsymbol{M}}^{\mp}(K)$. Then $\mathcal{M}^{\pm}$ are open, connected, globally hyperbolic subsets of*

$\boldsymbol{M}[\boldsymbol{h}]$ *for any* $\boldsymbol{h} \in H(\boldsymbol{M}; K)$*. Moreover,* $\boldsymbol{M}|_{\mathcal{M}^\pm} = \boldsymbol{M}[\boldsymbol{h}]|_{\mathcal{M}^\pm}$ *and the canonical inclusions* $\boldsymbol{M}[\boldsymbol{h}]|_{\mathcal{M}^\pm} \to \boldsymbol{M}[\boldsymbol{h}]$ *are Cauchy morphisms.*

Remark: We do not assume that K is connected.

Proof: As K is compact, $J^\pm_{\boldsymbol{M}}(K)$ are closed so $\mathcal{M}^\pm$ are open. We now claim that $J^\pm_{\boldsymbol{M}[\boldsymbol{h}]}(K) = J^\pm_{\boldsymbol{M}}(K)$ for any $h \in H(\boldsymbol{M}; K)$. To show this (for the (+) case), take any $q \in \mathcal{M} \setminus K$ with $q \in J^+_{\boldsymbol{M}[\boldsymbol{h}]}(K)$. Then there is a future-directed $\boldsymbol{M}[\boldsymbol{h}]$-causal curve $\gamma : [0,1] \to \mathcal{M}$ with $\gamma(0) \in K$ and $\gamma(1) = q$. Defining $\tau_* = \sup \gamma^{-1}(K)$ we have $\tau_* < 1$ and $\gamma(\tau_*) \in K$. As $\boldsymbol{h}$ is supported in K, the curve $\gamma|_{[\tau_*,1]}$ is also $\boldsymbol{M}[\boldsymbol{h}']$-causal for any $\boldsymbol{h}' \in H(\boldsymbol{M}; K)$, so $q \in J^+_{\boldsymbol{M}[\boldsymbol{h}']}(K)$. Thus $J^+_{\boldsymbol{M}[\boldsymbol{h}]}(K) \subset J^+_{\boldsymbol{M}[\boldsymbol{h}']}(K)$; reversing the roles of $\boldsymbol{h}$ and $\boldsymbol{h}'$ the two sets are therefore equal for arbitrary $\boldsymbol{h}, \boldsymbol{h}' \in H(\boldsymbol{M}; K)$. Setting $\boldsymbol{h}' = \boldsymbol{0}$ the (+)-case of the claim is established; the (−)-case is analogous.

To establish global hyperbolicity, take $p, q \in \mathcal{M}^-$ and γ a future-directed $\boldsymbol{M}[\boldsymbol{h}]$-causal curve from p to q. If γ leaves $\mathcal{M}^-$ then it contains a point of $J^+_{\boldsymbol{M}[\boldsymbol{h}]}(K)$; hence $q \in J^+_{\boldsymbol{M}[\boldsymbol{h}]}(K)$, which is a contradiction. Thus γ is contained within $\mathcal{M}$, as required.

Connectedness is proved as follows. Take any $p, q \in \mathcal{M}^-$; then we may find a $\boldsymbol{M}$-Cauchy surface Σ that is contained in $\mathcal{M}^-$ and lies to the past of both p, q.[20] As there are past-directed causal curves joining each of p and q to Σ, which is path-connected,[21] we conclude that p is path-connected to q. As p, q are arbitrary, we deduce that $\mathcal{M}^-$ is path-connected and hence connected.

Finally, as $\mathcal{M}^\pm$ contain $\boldsymbol{M}[\boldsymbol{h}]$-Cauchy surfaces (using a similar argument to that in footnote 20), the canonical inclusions of $\boldsymbol{M}[\boldsymbol{h}]|_{\mathcal{M}^\pm} \to \boldsymbol{M}[\boldsymbol{h}]$ are Cauchy morphisms. Moreover, $\boldsymbol{M}[\boldsymbol{h}]|_{\mathcal{M}^\pm} = (\mathcal{M}^\pm, \boldsymbol{g}|_{\mathcal{M}^\pm}, \mathfrak{o}|_{\mathcal{M}^\pm}, \mathfrak{t}|_{\mathcal{M}^\pm}) = \boldsymbol{M}|_{\mathcal{M}^\pm}$ for all $\boldsymbol{h} \in H(\boldsymbol{M})$. $\square$

Proof of Lemma 3.4: We prove the (+) case, thus supposing that the range of ψ is contained in $\mathcal{M} \setminus J^-_{\boldsymbol{M}}(K)$. Then we have $\psi^*\boldsymbol{h} = \boldsymbol{0}$ for $\boldsymbol{h} \in H(\boldsymbol{M}; K)$ and it follows straightforwardly that the underlying embedding of ψ induces $\psi[\boldsymbol{h}] : \boldsymbol{L} \to \boldsymbol{M}[\boldsymbol{h}]$. By Lem. 3.2 the set $\mathcal{M}^+ = \mathcal{M} \setminus J^-_{\boldsymbol{M}}(\operatorname{supp} \boldsymbol{h})$ is a globally hyperbolic subset of $\boldsymbol{M}$ and $\boldsymbol{M}[\boldsymbol{h}]$; as $\psi(\boldsymbol{L}) \subset \mathcal{M}^+ \subset \mathcal{M}$, the morphisms ψ and $\psi[\boldsymbol{h}]$ factor via the inclusion morphisms $\imath^+_{\boldsymbol{M}}[\boldsymbol{h}] : \boldsymbol{M}^+[\boldsymbol{h}] \to \boldsymbol{M}$ and $\jmath^+_{\boldsymbol{M}}[\boldsymbol{h}] : \boldsymbol{M}^+[\boldsymbol{h}] \to \boldsymbol{M}[\boldsymbol{h}]$ respectively, i.e.,

$$\psi = \imath^+_{\boldsymbol{M}}[\boldsymbol{h}] \circ \varphi^+, \qquad \psi[\boldsymbol{h}] = \jmath^+_{\boldsymbol{M}}[\boldsymbol{h}] \circ \varphi^+.$$

for $\varphi^+ : \boldsymbol{L} \to \boldsymbol{M}^+[\boldsymbol{h}]$.

If ψ is Cauchy then $\psi(\boldsymbol{L})$ contains a Cauchy surface for $\boldsymbol{M}$ and hence $\boldsymbol{M}^+[\boldsymbol{h}]$ (as $\psi(\boldsymbol{L}) \subset \mathcal{M}^+$). Thus φ^+ is Cauchy. As $\jmath^+_{\boldsymbol{M}}[\boldsymbol{h}]$ is Cauchy and the composite of Cauchy morphisms is Cauchy, it follows that $\psi[\boldsymbol{h}]$ is also Cauchy. $\square$

The next task is to prove that the push-forward of a globally hyperbolic perturbation under a Loc (or Loc_0) morphism is again a globally hyperbolic perturbation (Lemma A.7 below). This is broken into steps as follows.

[20] Let $\mathcal{T} : \mathcal{M} \to \mathbb{R}$ be a Cauchy temporal function for $\boldsymbol{M}$, which exists by [4, Thm 1.1]; then $\mathcal{T}(K)$ is compact and hence $\mathcal{T}(J^+_{\boldsymbol{M}}(K)) \subset [\tau, \infty)$ for some $\tau \in \mathbb{R}$, which, without loss of generality may be chosen so that $\tau < \min\{\mathcal{T}(p), \mathcal{T}(q), 0\}$. Then $\Sigma = \mathcal{T}^{-1}(\{2\tau\})$ meets the requirements; by choice of $\mathcal{T}$ we may additionally arrange that Σ be spacelike.

[21] It is connected [44, Prop. 14.31] and therefore path-connected, because it is a topological manifold.

Lemma A.5 *Suppose that K is a compact subset of a globally hyperbolic spacetime $\boldsymbol{M} = (\mathcal{M}, \boldsymbol{g}, \mathfrak{o}, \mathfrak{t}) \in \mathsf{Loc}$ and that $\gamma : I \to \mathcal{M}$ is an inextendible future-directed $\boldsymbol{M}$-timelike curve, where I is an open interval of $\mathbb{R}$. Then $\gamma^{-1}(K)$ is bounded.*

Proof: Choose a Cauchy temporal function $\mathcal{T}$ on $\boldsymbol{M}$; then $\mathcal{T}(K)$ is compact and contained in some interval (τ^-, τ^+). Then $\Sigma^\pm = \mathcal{T}^{-1}(\tau^\pm)$ are Cauchy surfaces of $\boldsymbol{M}$ to the past $(-)$ and future $(+)$ of K, i.e., $J^\pm_{\boldsymbol{M}}(\Sigma^\pm) \cap K = \emptyset$. As it is inextendible, γ intersects $\Sigma^\pm$ at unique $t^\pm \in I$ and it is clear that $\gamma^{-1}(K) \subset (t^+, t^-)$ because $\gamma(t)$ lies in $J^-_{\boldsymbol{M}}(\Sigma^-)$ for $t < t^-$ (resp., $J^+_{\boldsymbol{M}}(\Sigma^+)$ for $t > t^+$) and does not intersect K in this interval. $\square$

Lemma A.6 *Suppose $\boldsymbol{M} = (\mathcal{M}, \boldsymbol{g}, \mathfrak{o}, \mathfrak{t}) \in \mathsf{Loc}$ and let K be a compact subset of $\mathcal{M}$ contained in an open $\boldsymbol{M}$-causally convex subset U that has at most finitely many connected components. Let Σ be a Cauchy surface of $\boldsymbol{M}$ to the past of K, i.e., $K \subset I^+_{\boldsymbol{M}}(\Sigma)$. Suppose $\boldsymbol{g}'$ is a time-orientable Lorentz metric on $\mathcal{M}$ with time-orientation $\mathfrak{t}'$ such that $\boldsymbol{g}' = \boldsymbol{g}$, $\mathfrak{t}' = \mathfrak{t}$ outside K, and so that $\boldsymbol{U} = (U, \boldsymbol{g}'|_U, \mathfrak{o}|_U, \mathfrak{t}'|_U) \in \mathsf{Loc}$. Then:*

- *(i) if γ is a $(\boldsymbol{g}', \mathfrak{t}')$-causal curve in $\mathcal{M}$ with endpoints in U then γ is contained in U;*
- *(ii) if $\gamma : \mathbb{R} \to \mathcal{M}$ is an inextendible $(\boldsymbol{g}', \mathfrak{t}')$-timelike curve intersecting K then $\gamma^{-1}(U)$ is an open interval and $\gamma^{-1}(K)$ is bounded;*
- *(iii) any inextendible $\boldsymbol{g}'$-timelike curve $\gamma : \mathbb{R} \to \mathcal{M}$ intersects Σ exactly once;*
- *(iv) the spacetime $\boldsymbol{M}' = (\mathcal{M}, \boldsymbol{g}', \mathfrak{o}, \mathfrak{t}')$ is globally hyperbolic, i.e., $\boldsymbol{M}' \in \mathsf{Loc}$.*

Proof: (i) Suppose $\gamma : [0,1] \to \mathcal{M}$ is $(\boldsymbol{g}', \mathfrak{t}')$-causal with $\gamma(0), \gamma(1) \in U$, but $\gamma(t) \notin U$ for some $t \in (0,1)$. Then there are t_0, t_1 with $0 < t_0 < t < t_1 < 1$ such that $\gamma(t_0), \gamma(t_1) \in U$ but $\gamma|_{[t_0,t_1]}$ does not intersect K. Hence $\gamma|_{[t_0,t_1]}$ is $\boldsymbol{M}$-causal and therefore contained in U by causal convexity. This is a contradiction.
(ii) An immediate corollary of (i) is that $I = \gamma^{-1}(U)$ is an open convex subset of $\mathbb{R}$, i.e., an open interval. Now the restriction of γ to I is an inextendible future-directed timelike curve in the globally hyperbolic spacetime $\boldsymbol{U}$. Applying Lemma A.5, we find that $(\gamma|_I)^{-1}(K) = \gamma^{-1}(K)$ is bounded.
(iii) If γ does not intersect the interior of K, it is also $\boldsymbol{M}$-timelike and therefore intersects Σ exactly once. If γ does intersect $\mathrm{int}(K) \subset U$, then $\gamma^{-1}(K)$ is bounded from below by (ii) and $t_0 = \inf \gamma^{-1}(K)$ is finite. As any portion of γ outside K is $\boldsymbol{M}$-timelike and future-directed, we have

$$\gamma(t) \in J^+_{\boldsymbol{M}}(\gamma(\sup\{t' \le t : \ \gamma(t') \in K\})) \subset J^+_{\boldsymbol{M}}(K)$$

for any $t > t_0$. Thus $\gamma|_{(t_0,\infty)}$ does not intersect Σ, while the past-inextendible portion $\gamma|_{(-\infty,t_0]}$ intersects Σ exactly once because $\gamma(t_0) \in K$ lies to the future of Σ.
(iv) It follows immediately from (iii) that Σ is a Cauchy surface for the spacetime $(\mathcal{M}, \boldsymbol{g}', \mathfrak{o}', \mathfrak{t}')$, which is therefore globally hyperbolic. $\square$

We can now prove the covariance property of globally hyperbolic perturbations.

Lemma A.7 *Suppose $\psi : \boldsymbol{L} \to \boldsymbol{M}$ in Loc. Then $\psi_*(H(\boldsymbol{L})) \subset H(\boldsymbol{M})$. (In particular, this applies to all morphisms in Loc_0.)*

Proof: Write $\boldsymbol{M} = (\mathcal{M}, \boldsymbol{g}, \mathfrak{o}, \mathfrak{t})$, $\boldsymbol{L} = (\mathcal{L}, \psi^*\boldsymbol{g}, \psi^*\mathfrak{o}, \psi^*\mathfrak{t})$ and $U = \psi(\mathcal{L})$, $K = \psi(\operatorname{supp} \boldsymbol{h})$, where $\boldsymbol{h} \in H(\boldsymbol{L})$. Then $\boldsymbol{g}' = \boldsymbol{g} + \psi_*\boldsymbol{h}$ is a Lorentz metric on $\mathcal{M}$. To show that it is time-orientable, let T_1 (resp., T_2) be a $\boldsymbol{L}[\boldsymbol{h}]$-timelike (resp., $\boldsymbol{M}$-timelike) nowhere zero, future-pointing vector field on $\mathcal{L}$ (resp., $\mathcal{M}$). Let $\chi \in C_0^\infty(\mathcal{M})$ be nonnegative, with $\chi = 1$ on K and $\chi = 0$ outside U. Then $\chi\psi_*T_1 + (1-\chi)T_2$ is nowhere zero and $\boldsymbol{g}'$-timelike, and therefore defines a time-orientation $\mathfrak{t}'$ of $\boldsymbol{g}'$ that agrees with $\mathfrak{t}$ outside $\psi(K)$. As there exist $\boldsymbol{M}$-Cauchy surfaces to the past of K, Lem. A.6(iv) entails that $(\mathcal{M}, \boldsymbol{g}', \mathfrak{o}, \mathfrak{t}') \in \mathsf{Loc}$, i.e., $\psi_*\boldsymbol{h} \in H(\boldsymbol{N})$. □

A.3 Causal complements and (multi-)diamonds

Finally, we give a number of results relating to causal structure and multi-diamonds in globally hyperbolic spacetimes. Similar results appear elsewhere (e.g., [12, Appx B], [11, §2], [46, §3]) but we are not aware of a full presentation of all the results needed in the body of this paper. Notation and terminology varies in the literature and the definition of causal complement is not always made clear (the cited references are exceptions to this). It is hoped that this appendix may be useful more widely.

Recall that we have two notions of causal complement in a globally hyperbolic spacetime $\boldsymbol{M}$: $O^\perp = \mathcal{M} \setminus J_{\boldsymbol{M}}(O)$ and $O' = \mathcal{M} \setminus \mathrm{cl}(J_{\boldsymbol{M}}(O))$. Clearly O' is always an open set.

Lemma A.8 *Let O be an open subset of a globally hyperbolic spacetime $\boldsymbol{M}$. Then $J^\pm_{\boldsymbol{M}}(O)$ are open, and $O \subset O''$.*

Proof: If $q \in J^\pm_{\boldsymbol{M}}(O)$ then O has nontrivial intersection with the (closed) set $J^\mp_{\boldsymbol{M}}(q)$, which is the closure of $I^\mp_{\boldsymbol{M}}(q)$ as $\boldsymbol{M}$ is globally hyperbolic ([44], Lem. 14.6). Thus O intersects $I^\mp_{\boldsymbol{M}}(q)$, so $q \in I^\pm_{\boldsymbol{M}}(p')$ for some $p' \in O$, and we have shown that $J^\pm_{\boldsymbol{M}}(O) \subset I^\pm_{\boldsymbol{M}}(O) = \mathrm{int}(J^\pm_{\boldsymbol{M}}(O))$. As $O' \cap J_{\boldsymbol{M}}(O)$ is empty, so is $J_{\boldsymbol{M}}(O') \cap O$ and hence $\mathrm{cl}(J_{\boldsymbol{M}}(O')) \cap O$. Thus $O \subset O''$. □

Lemma A.9 *Let Σ be an acausal Cauchy surface in globally hyperbolic spacetime $\boldsymbol{M}$ and let S be an open subset of Σ such that $\mathrm{cl}\, S$ has nontrivial complement in Σ. Then $S'' = D_{\boldsymbol{M}}(S) = D_{\boldsymbol{M}}(S)''$. In particular, every multi-diamond O is causally complete in the sense that $O = O''$.*

Proof: First, using $D_{\boldsymbol{M}}(S) \subset J_{\boldsymbol{M}}(S)$, observe that

$$D_{\boldsymbol{M}}(S)' = \mathcal{M} \setminus \mathrm{cl}\, J_{\boldsymbol{M}}(D_{\boldsymbol{M}}(S)) = \mathcal{M} \setminus \mathrm{cl}\, J_{\boldsymbol{M}}(S) = S'$$

and hence $S'' = D_{\boldsymbol{M}}(S)''$ (this holds for any subset S of $\boldsymbol{M}$; similarly, we also have $D_{\boldsymbol{M}}(S)^\perp = S^\perp$ and thus $S^{\perp\perp} = D_{\boldsymbol{M}}(S)^{\perp\perp}$ for any subset S); it remains to show that $D_{\boldsymbol{M}}(S)$ is causally complete.

As S is open in Σ, it inherits the property of being an acausal topological hypersurface [44, 14.24] from Σ; accordingly $O = D_{\boldsymbol{M}}(S)$ is an open subset of $\boldsymbol{M}$ [44, 14.42]. We observe that $\Sigma \setminus \mathrm{cl}(S) \subset O'$; if not, then we may find $q \in \Sigma \setminus \mathrm{cl}(S)$ and $q_n \to q$ with $q_n \in J_{\boldsymbol{M}}(O)$. Choose an open neighbourhood U of q in Σ that does not intersect $\mathrm{cl}(S)$,

then $D_{\boldsymbol{M}}(U)$ is an open neighbourhood of q that contains q_n for sufficiently large n. But $q_n \in J_{\boldsymbol{M}}(O) = J_{\boldsymbol{M}}(S)$ contradicts $q_n \in D_{\boldsymbol{M}}(U)$.

To establish causal completeness it suffices to show $O'' \subset O$. If $p \notin O$ there is an inextendible causal curve through p intersecting Σ at $q \notin S$. Assume without loss that $p \in J^+_{\boldsymbol{M}}(q)$. Then there are points $p_n \to p$ with $p_n \in I^+_{\boldsymbol{M}}(q)$ and hence neighbourhoods U_n of q with $U_n \subset I^-_{\boldsymbol{M}}(p_n)$. Each U_n must intersect $\Sigma \setminus \mathrm{cl}(S)$ nontrivially, so $p_n \in J^+_{\boldsymbol{M}}(\Sigma \setminus \mathrm{cl}(S)) \subset J_{\boldsymbol{M}}(O')$. Hence $p \in \mathrm{cl}(J_{\boldsymbol{M}}(O'))$ i.e., $p \notin O''$. Thus $O'' \subset O$, so $O'' = O$.

If O is a multi-diamond then $O = D_{\boldsymbol{M}}(S)$ where S meets the above hypotheses; hence $O = O''$. $\square$

Lemma A.10 *Suppose K is a compact subset of globally hyperbolic spacetime $\boldsymbol{M}$. Then (i) $K^\perp$ is open, $K^{\perp\perp}$ is closed, and $K \subset K^{\perp\perp}$; (ii) $K^{\perp\perp\perp} = K^\perp$ and $K^{\perp\perp}$ is causally complete with respect to $\perp$; (iii) $K^{\perp\perp}$ is causally convex. If, in addition, K has a multi-diamond neighbourhood O then $K^{\perp\perp}$ is compact and contained in* $\mathrm{cl}(O)$*.*

Remark: In general $K^{\perp\perp}$ need not be compact, e.g., if K contains a Cauchy surface for $\boldsymbol{M}$. As another example, let K be a closed ball of radius 1 in the $t = 0$ plane of the $|t| < 1/2$ portion of Minkowski space in standard coordinates; then $K^{\perp\perp}$ is the $|t| < 1/2$ portion of the diamond based on the interior of K, and is noncompact.
Proof: (i) As K is compact, $J_{\boldsymbol{M}}(K)$ is closed and $K^\perp = \mathcal{M} \setminus J_{\boldsymbol{M}}(K)$ is therefore open. Hence, by Lem. A.8, $J_{\boldsymbol{M}}(K^\perp)$ is open and $K^{\perp\perp}$ is closed. Moreover, as $K^\perp \cap J_{\boldsymbol{M}}(K)$ is empty, so is $J_{\boldsymbol{M}}(K^\perp) \cap K$; hence we see that $K \subset K^{\perp\perp}$.
(ii) If $p \in K^{\perp\perp}$ then $J_{\boldsymbol{M}}(p) \subset J_{\boldsymbol{M}}(K)$; otherwise, $J_{\boldsymbol{M}}(p)$ would intersect $K^\perp$, giving $p \in J_{\boldsymbol{M}}(K^\perp)$ and a contradiction. Thus $J_{\boldsymbol{M}}(K^{\perp\perp}) = J_{\boldsymbol{M}}(K)$ and so $K^{\perp\perp\perp} = K^\perp$. In particular, $K^{\perp\perp}$ is casually complete with respect to $\perp$.
(iii) Take any $p, q \in K^{\perp\perp}$. If a future-directed causal curve γ joins p and q but leaves $K^{\perp\perp}$ there must be $r \in J_{\boldsymbol{M}}(K^\perp)$ such that $q \in J^+_{\boldsymbol{M}}(r)$, $p \in J^-_{\boldsymbol{M}}(r)$. Thus one or both of p, q belong to $J_{\boldsymbol{M}}(K^\perp)$, which is a contradiction. Hence $K^{\perp\perp}$ is causally convex and therefore a closed globally hyperbolic subset of $\boldsymbol{M}$.
Finally, if K has a multi-diamond neighbourhood O, then $O' \subset K^\perp$, and hence $J_{\boldsymbol{M}}(O') \subset J_{\boldsymbol{M}}(K^\perp)$. Hence $K^{\perp\perp} \subset \mathcal{M} \setminus J_{\boldsymbol{M}}(O') = \mathrm{cl}(O'') = \mathrm{cl}(O)$, which is compact. Accordingly, $K^{\perp\perp}$ is a closed subset contained in a compact set, and hence compact. $\square$

Lemma A.11 *Let $\boldsymbol{M}$ be a globally hyperbolic spacetime. (i) Suppose O_1 and O_2 are open subsets of $\boldsymbol{M}$, with O_1 relatively compact and $\mathrm{cl}(O_1) \subset O_2$. Then $\mathrm{cl}(J_{\boldsymbol{M}}(O_1)) \subset J_{\boldsymbol{M}}(O_2)$. (ii) Suppose O_n ($n \in \mathbb{N}$) is a sequence of relatively compact subsets of $\boldsymbol{M}$ with $\mathrm{cl}(O_{n+1}) \subset O_n$ for all $n \in \mathbb{N}$ and $\bigcap_{n\in\mathbb{N}} O_n = K$ compact. Then*

$$J_{\boldsymbol{M}}(K) = \bigcap_{n\in\mathbb{N}} J_{\boldsymbol{M}}(O_n) = \bigcap_{n\in\mathbb{N}} \mathrm{cl}(J_{\boldsymbol{M}}(O_n)), \quad \textit{and hence} \quad K^\perp = \bigcup_{n\in\mathbb{N}} O_n'.$$

Proof: (i) We calculate

$$\mathrm{cl}(J_{\boldsymbol{M}}(O_1)) \subset \mathrm{cl}(J_{\boldsymbol{M}}(\mathrm{cl}(O_1))) = J_{\boldsymbol{M}}(\mathrm{cl}(O_1)) \subset J_{\boldsymbol{M}}(O_2)$$

using the fact that $J_{\boldsymbol{M}}(\mathrm{cl}(O_1))$ is closed.
(ii) The inclusion $J_{\boldsymbol{M}}(K) \subset \bigcap_{n\in\mathbb{N}} J_{\boldsymbol{M}}(O_n)$ is immediate from $K \subset O_n$ for all n. On the other hand, if $p \in \bigcap_{n\in\mathbb{N}} J_{\boldsymbol{M}}(O_n)$ then there exist $q_n \in J_{\boldsymbol{M}}(p) \cap O_n$ for all n. As all q_n are contained in the relatively compact set O_1 we may pass to a convergent subsequence q_{n_r} with limit $q \in \mathrm{cl}(O_1)$; as all but finitely many of the q_{n_r} are contained in each O_{m+1} ($m = 1, 2, \ldots$), we also have $q \in \mathrm{cl}(O_{m+1}) \subset O_m$ for each $m \in \mathbb{N}$ and hence $q \in K$. As the q_{n_r} lie in the closed set $J_{\boldsymbol{M}}(p)$, we additionally have $q \in K \cap J_{\boldsymbol{M}}(p)$ and hence conclude that $p \in J_{\boldsymbol{M}}(K)$. Accordingly we have proved the first of the required equalities. By part (i) we have $\mathrm{cl}(J_{\boldsymbol{M}}(O_{n+1})) \subset J_{\boldsymbol{M}}(O_n)$ for all n from which the second equality follows. Taking complements in $\boldsymbol{M}$ we obtain the required formula for $K^\perp$. $\square$

Lemma A.12 *Let S be a subset of a time-oriented Lorentzian spacetime $\boldsymbol{M}$ such that $J^+_{\boldsymbol{M}}(\mathrm{cl}(S))$ is closed (for example, if S is a relatively compact subset of a globally hyperbolic spacetime). Then*

$$J^+_{\boldsymbol{M}}(\mathrm{cl}(S)) = \mathrm{cl}(I^+_{\boldsymbol{M}}(S)) = \mathrm{cl}(J^+_{\boldsymbol{M}}(S)). \tag{A.1}$$

The analogous result holds for causal and chronological pasts. If both $J^\pm_{\boldsymbol{M}}(\mathrm{cl}(S))$ are closed then $J_{\boldsymbol{M}}(\mathrm{cl}(S)) = \mathrm{cl}(J_{\boldsymbol{M}}(S))$ and hence $(\mathrm{cl}(S))^\perp = S'$.

Proof: Owing to the hypothesis, we have

$$J^+_{\boldsymbol{M}}(\mathrm{cl}(S)) = \mathrm{cl}(I^+_{\boldsymbol{M}}(\mathrm{cl}(S))) = \mathrm{cl}(I^+_{\boldsymbol{M}}(S)) \subset \mathrm{cl}(J^+_{\boldsymbol{M}}(S)) \subset \mathrm{cl}(J^+_{\boldsymbol{M}}(\mathrm{cl}(S))) = J^+_{\boldsymbol{M}}(\mathrm{cl}(S))$$

using the standard results Lemma 14.6(2) in [44] and Prop. 2.11 in [45] for the first two equalities. This establishes Eq. (A.1); the remaining statements are trivial. $\square$

Lemma A.13 *Let S be any subset in a globally hyperbolic spacetime $\boldsymbol{M}$. Then the Cauchy development obeys $D_{\boldsymbol{M}}(S) \subset S^{\perp\perp}$, with equality if S lies in an acausal Cauchy surface of $\boldsymbol{M}$.*

Remark: The example of $S = \{p, q\}$ for $q \in J^+_{\boldsymbol{M}}(p)$, for which $D_{\boldsymbol{M}}(S) = S$, $S^{\perp\perp} = J^+_{\boldsymbol{M}}(p) \cap J^-_{\boldsymbol{M}}(q)$, shows that equality cannot be expected in general.
Proof: If $p \in D_{\boldsymbol{M}}(S)$ then every inextendible causal curve through p intersects S. Thus any point causally connected to p is causally connected to S, i.e., $p \notin J_{\boldsymbol{M}}(S^\perp)$ and hence $p \in S^{\perp\perp}$. If $S \subset \Sigma$, an acausal Cauchy surface of $\boldsymbol{M}$, then $\Sigma \setminus S \subset S^\perp$. Accordingly, any inextendible causal curve through $p \in S^{\perp\perp}$ must cut Σ in S, so $S^{\perp\perp} = D_{\boldsymbol{M}}(S)$ in this case. $\square$

Lemma A.14 *Suppose D is a multi-diamond, with base B in spacetime $\boldsymbol{M} \in \mathsf{Loc}$. If K is any compact subset of D then $K \subset \tilde{K}^{\perp\perp}$ for a compact subset $\tilde{K}$ of B (hence $\tilde{K} \in \mathscr{K}(\boldsymbol{M}; D)$).*

Proof: Suppose Σ is a spacelike Cauchy surface for $\boldsymbol{M}$ with $B \subset \Sigma$. Then $J_{\boldsymbol{M}}(K) \cap \Sigma$ is compact and contained in B, which has a finite number R of connected components B_r. Each B_r is contained in a chart (U_r, ϕ_r) of Σ in which $\phi_r(B_r)$ is an open ball; we may

choose a compact set K_r so that $\phi_r(K_r)$ is the closure of a slightly smaller ball with the same centre and so that K_r contains $J_{\boldsymbol{M}}(K)\cap\Sigma\cap B_r$. Then $\tilde{K}=\bigcup_{r=1}^{R} K_r$ is compact and contains $J_{\boldsymbol{M}}(K)\cap\Sigma$. Moreover, $K\subset D_{\boldsymbol{M}}(\tilde{K})=\tilde{K}^{\perp\perp}$ by Lemma A.13 and the fact that spacelike Cauchy surfaces are acausal [44, Lem. 14.42]. Finally, $\tilde{K}$ has a multi-diamond neighbourhood D, with base $B\subset D$, so $\tilde{K}\in\mathscr{K}(\boldsymbol{M};D)$. $\square$

Lemma A.15 *If $\psi:\boldsymbol{M}\to\boldsymbol{N}$ in* Loc *then $\psi(K^{\perp\perp})=\psi(K)^{\perp\perp}$ for all $K\in\mathscr{K}(\boldsymbol{M})$.*

Proof: Observe first that for any subset $S\subset\boldsymbol{M}$, we have $J_{\boldsymbol{N}}(\psi(S))\cap\psi(\boldsymbol{M})=\psi(J_{\boldsymbol{M}}(S))$ by causal convexity of $\psi(\boldsymbol{M})$ and hence $\psi(S)^{\perp}\cap\psi(\boldsymbol{M})=\psi(S^{\perp})$, using also the injectivity of ψ. It follows that

$$\psi(K^{\perp\perp})=\psi(K^{\perp})^{\perp}\cap\psi(\boldsymbol{M})=(\psi(K)^{\perp}\cap\psi(\boldsymbol{M}))^{\perp}\cap\psi(\boldsymbol{M}).$$

But as K has a multi-diamond neighbourhood D in $\boldsymbol{M}$, $K^{\perp\perp}\subset\mathrm{cl}(D)$ (Lem. A.10); similarly, as $\psi(D)$ is a multi-diamond in $\boldsymbol{N}$ we have $\psi(K)^{\perp\perp}\subset\mathrm{cl}(\psi(D))\subset\psi(\boldsymbol{M})$ and hence

$$\psi(K^{\perp\perp})=(\psi(K)^{\perp}\cap\psi(\boldsymbol{M}))^{\perp}\supset\psi(K)^{\perp\perp}.$$

Now take any point $p\in K^{\perp\perp}$ and suppose for a contradiction that $\psi(p)\in\boldsymbol{N}\setminus\psi(K)^{\perp\perp}=J_{\boldsymbol{N}}(\psi(K)^{\perp})$. By causal convexity of $\psi(\boldsymbol{M})$, $\psi(p)$ would lie in $J_{\boldsymbol{N}}(\psi(K)^{\perp})$ only if $p\in J_{\boldsymbol{M}}(K^{\perp})$, which would contradict the assumption that $p\in K^{\perp\perp}$. Accordingly, we have $\psi(p)\subset J_{\boldsymbol{N}}(q)$ for some $q\in\psi(K)^{\perp}\setminus\psi(K^{\perp})$, which must therefore lie outside $\psi(\boldsymbol{M})$ because $\psi(K^{\perp})=\psi(K)^{\perp}\cap\psi(\boldsymbol{M})$ as shown above. Without loss of generality, we may suppose that $\psi(p)$ lies to the future of q along smooth causal curve γ. The pre-image of γ under ψ is a connected future-directed smooth causal curve, which is past-inextendible in $\boldsymbol{M}$ and therefore contains points outside $J^{+}_{\boldsymbol{M}}(K)$. Take any such point r; r cannot lie in $J^{-}_{\boldsymbol{M}}(K)$ (otherwise $q\in J^{-}_{\boldsymbol{N}}(\psi(K))$) and hence $r\in K^{\perp}$. But this entails that $p\in J_{\boldsymbol{M}}(K^{\perp})$, contradicting the initial assumption $p\in K^{\perp\perp}$. $\square$

B Subobjects, intersections and unions

We summarise the basic properties of subobjects that are used in the body of the text. For completeness, we also include some standard definitions of category theory (although we take the basic definition of a category for granted). To a large extent we follow [21].

In a general category C, then, a morphism f is described as *monic* (or as a *monomorphism*) iff it is left-cancellable, so $f\circ g=f\circ h$ implies $g=h$, and as *epic* (or as an *epimorphism*) iff it is right-cancellable, so $g\circ f=h\circ f$ implies $g=h$. An object $\mho$ of C is *initial* if there is a unique morphism $\mho_A:\mho\to A$ for each object A of C. A monic will be equivalently described as defining a *subobject* of its codomain, so that $m:M\to A$ is a subobject of A. In cases where the morphism $\mho_A$ is monic, we will describe this as the trivial subobject of A. Subobjects $M\stackrel{m}{\to}A$ and $M'\stackrel{m'}{\to}A$ are isomorphic iff there exists an isomorphism $f:M\to M'$ such that $m=m'\circ f$, in which case we write $m\cong m'$; in the

case where $m = m' \circ f$ for some f that is not necessarily an isomorphism, we write $m \leq m'$ (f is uniquely specified because m' is monic).

A category C has *equalizers* if it satisfies the following condition: for every pair of morphisms $f, g : A \to B$ there is a morphism h such that $f \circ h = g \circ h$ and such that if k is any morphism such that $f \circ k = g \circ k$ then k factorizes uniquely via h, i.e., $k = h \circ m$ for a unique morphism m; h is said to be an equalizer of f and g in this situation.

Given a collection $(m_i)_{i \in I}$ [in which I is a class] of subobjects of A their intersection and union may be defined as follows: An *intersection* is a subobject $m : M \to A$ with the following properties:

1. m factorises via each m_i as $m = m_i \circ j_i$;
2. given any $f : B \to A$ factorising via each m_i as $f = m_i \circ k_i$, there exists a unique $g : B \to M$ such that $j_i \circ g = k_i$ for all $i \in I$, and hence $f = m \circ g$.

These properties define m up to isomorphism and we write

$$m \cong \bigwedge_{i \in I} m_i : \bigwedge_{i \in I} M_i \to A.$$

The category C is said to have intersections (with respect to monics) if every such collection of subjobjects has an intersection. More generally, one can define intersections with respect to a subclass $\mathcal{M}$ of monics [21].

Lemma B.1 *(a) With the above notation, if $(v_i)_{i \in I}$ are isomorphisms pre-composable with the (m_i) then $(m_i)_{i \in I}$ has an intersection if and only if $(m_i \circ v_i)_{i \in I}$ does, and*

$$\bigwedge_{i \in I} m_i \circ v_i \cong \bigwedge_{i \in I} m_i$$

(b) If $k : A \to A'$ is monic then $(k \circ m_i)_{i \in I}$ has an intersection if and only if $(m_i)_{i \in I}$ does; provided that I is nonempty[22] *we have*

$$k \circ \bigwedge_{i \in I} m_i \cong \bigwedge_{i \in I} k \circ m_i$$

Proof: (a) Suppose (m_i) has an intersection m with factorizations $m = m_i \circ j_i$. Then m also factorizes as $m = m_i \circ v_i \circ j'_i$ for $j'_i = v_i^{-1} \circ j_i$ and we will show that this defines an intersection of $(m_i \circ v_i)_{i \in I}$. Suppose f factorizes as $f = m_i \circ v_i \circ k_i$, then the intersection property of the (m_i) implies that there is a unique g such that $v_i \circ k_i = j_i \circ g$ and hence $k_i = j'_i \circ g$ for all $i \in I$. Thus $(m_i \circ v_i)_{i \in I}$ has m as an intersection. The reverse implication also follows from this argument.
(b) Suppose (m_i) has an intersection m with factorizations $m = m_i \circ j_i$; we must show that $k \circ m$ is an intersection of the $k \circ m_i$, with factorizations $k \circ m = (k \circ m_i) \circ j_i$. To this end, suppose there are factorizations $f = k \circ m_i \circ l_i$ for all i. As k is monic, this implies the

[22] The intersection of an empty class of subobjects of A is id_A.

existence of h such that $m_i \circ l_i = h$ for all i and (because $\bigwedge_i m_i$ exists), the existence of a unique g with $l_i = j_i \circ g$ for all i, which was to be shown. On the other hand, suppose that $(k \circ m_i)_{i \in I}$ have an intersection $h = k \circ m_i \circ j_i$. Again, as k is monic, we may write $h = k \circ m$ with $m = m_i \circ j_i$ for all i. To see that this defines an intersection of $(m_i)_{i \in I}$, suppose $f = m_i \circ l_i$ for all i. Then $k \circ f = k \circ m_i \circ l_i$ and (because $\bigwedge_i k \circ m_i$ exists) there is a unique g such that $l_i = j_i \circ g$, which was to be shown. □

On the other hand, the *union* is a subobject $m : M \to A$ with the following properties

1. every m_i factorises as $m_i = m \circ \tilde{m}_i$ (in which $\tilde{m}_i : M_i \to M$)
2. given any $f : A \to B$, if there exists a subobject $n : N \to B$ such that every $f \circ m_i$ factorises as $n \circ \tilde{n}_i$, then there is a unique morphism $\tilde{f} : M \to N$ such that $n \circ \tilde{f} = f \circ m$ and $\tilde{f} \circ \tilde{m}_i = \tilde{n}_i$ for all $i \in I$.

Property (2) can be displayed diagrammatically as the commuting diagram

$$\begin{tikzcd} & M_i \arrow[dl, "\tilde{m}_i"'] \arrow[dr, "\tilde{n}_i"] & \\ M \arrow[rr, "\tilde{f}"] \arrow[d, "m"'] & & N \arrow[d, "n"] \\ A \arrow[rr, "f"] & & B \end{tikzcd} \tag{B.1}$$

(in which it is tacit that $m_i = m \circ \tilde{m}_i$).

It is easy to see that this defines the union subobject up to isomorphism; we therefore write [following [21] §1.9]

$$m \cong \bigvee_{i \in I} m_i : \bigvee_{i \in I} M_i \to A$$

The union always exists if C has intersections and also has pull-backs with respect to monics in the following sense: whenever $f : X \to Y$ and $n : N \to Y$ is a subobject, there is a subobject $m : M \to X$ and a morphism $f' : M \to N$ such that $n \circ f' = f \circ m$, and if there are morphisms g and h such that $n \circ h = f \circ g$ then there is a unique t such that $m \circ t = g$, whereupon also $h = f' \circ t$.

Lemma B.2 *Let $(m_i)_{i \in I}$ (resp., $(n_j)_{j \in J}$) be a class-indexed family of subobjects of $A \in \mathsf{C}$ with union $m : M \to A$ (resp., $n : N \to A$). If, to each $i \in I$ there is $j(i) \in J$ such that $m_i = n_{j(i)} \circ \mu_i$ for some μ_i, then there is a unique $\xi : M \to N$ such that $n \circ \xi = m$. If, additionally, $J \subset I$ and $n_j \cong m_j$ for each $j \in J$ then ξ is an isomorphism.*

Proof: Let $n_j = n \circ \hat{n}_j$ be the factorizations associated with $\bigvee_{j\in J} n_j$, and consider diagram (B.1), with $B = A$, $f = \mathrm{id}_A$ and $\tilde{n}_i = \hat{n}_{j(i)} \circ \mu_i$. As the outer portion commutes we deduce the existence of a unique ξ (replacing $\tilde{f}$) with the property stated. In the special case, we may apply this result again with the roles of m_i and n_j reversed, giving a unique η such that $m \circ \eta = n$. As m and n are monic, it follows that η and ξ are mutual inverses, hence isomorphisms. □

A useful consequence is that if I is a class and for each $i \in I$ there is a nonempty class J_i labelling subobjects m_{ij}, then we have the 'Fubini property'

$$\bigvee_{i\in I}\bigvee_{j\in J_i} m_{ij} \cong \bigvee_{(i,j)\in K} m_{ij} \cong \bigvee_{j\in J}\bigvee_{i\in I_j} m_{ij} \tag{B.2}$$

where $J = \bigcup_{i\in I} J_i$, $K = \{(i,j) \in I \times J : j \in J_i\}$ and $I_j = \{i \in I : j \in J_i\}$ $(j \in J)$.

Lemma B.3 *Suppose a category* C *has equalizers, and intersections and pullbacks with respect to monics. Let* $(m_i)_{i\in I}$ *be a class-indexed family of subobjects of* $A \in \mathsf{C}$ *with union* $m : M \to A$. *If* $h : A \to A$ *obeys* $h \circ m_i = m_i$ *for all* $i \in I$ *then* $h \circ m = m$.

Proof: We have $h \circ m_i = \mathrm{id}_A \circ m_i$ and hence a factorisation $m_i = g \circ \tilde{g}_i$ for each $i \in I$ where g is an equalizer of h and id_A (and is necessarily monic). In conjunction with the factorisation $m_i = m \circ \tilde{m}_i$ this induces a factorisation $m_i = n \circ \tilde{n}_i$ via the intersection (=pullback) $n : N \to A$ of g and m, corresponding to $n = g \circ k = m \circ \ell$. The outer portion of the diagram (B.1) commutes for all $i \in I$, with $B = A$, $f = \mathrm{id}_A$, and there is therefore a morphism $\tilde{f}$ to make the diagram commute in full. Consequently, $h \circ m = h \circ n \circ \tilde{f} = h \circ g \circ k \circ \tilde{f} = g \circ k \circ \tilde{f} = n \circ \tilde{f} = m$ as required. □

As we study categories in which all morphisms are monic, the existence of pull-backs with respect to monics follows from existence of intersections.

References

[1] Adámek, J., Herrlich, H., Strecker, G.E.: Abstract and concrete categories: the joy of cats. Repr. Theory Appl. Categ. pp. 1–507 (2006), reprint of the 1990 original [Wiley, New York]

[2] Beem, J.K., Ehrlich, P.E., Easley, K.L.: Global Lorentzian geometry, *Monographs and Textbooks in Pure and Applied Mathematics*, vol. 202, second edn. Marcel Dekker Inc., New York (1996)

[3] Bernal, A.N., Sánchez, M.: On smooth Cauchy hypersurfaces and Geroch's splitting theorem. Commun. Math. Phys. **243**, 461–470 (2003)

[4] Bernal, A.N., Sánchez, M.: Smoothness of time functions and the metric splitting of globally hyperbolic spacetimes. Commun. Math. Phys. **257**, 43–50 (2005), gr-qc/0401112

[5] Bernal, A.N., Sánchez, M.: Further results on the smoothability of Cauchy hypersurfaces and Cauchy time functions. Lett. Math. Phys. **77**, 183–197 (2006), gr-qc/0512095

[6] Bernal, A.N., Sánchez, M.: Globally hyperbolic spacetimes can be defined as causal instead of strongly causal. Class. Quantum Grav. **24**, 745–750 (2007), gr-qc/0611138

[7] Brunetti, R., Fredenhagen, K.: Microlocal analysis and interacting quantum field theories: Renormalization on physical backgrounds. Commun. Math. Phys. **208**, 623–661 (2000)

[8] Brunetti, R., Fredenhagen, K.: Quantum field theory on curved backgrounds. In: Quantum field theory on curved spacetimes, *Lecture Notes in Phys.*, vol. 786, pp. 129–155. Springer, Berlin (2009)

[9] Brunetti, R., Fredenhagen, K., Verch, R.: The generally covariant locality principle: A new paradigm for local quantum physics. Commun. Math. Phys. **237**, 31–68 (2003)

[10] Brunetti, R., Guido, D., Longo, R.: Modular localization and Wigner particles. Rev. Math. Phys. **14**, 759–785 (2002)

[11] Brunetti, R., Ruzzi, G.: Superselection sectors and general covariance. I. Commun. Math. Phys. **270**, 69–108 (2007)

[12] Brunetti, R., Ruzzi, G.: Quantum charges and spacetime topology: The emergence of new superselection sectors. Commun. Math. Phys. **287**, 523–563 (2009)

[13] Buchholz, D., Dreyer, O., Florig, M., Summers, S.J.: Geometric modular action and spacetime symmetry groups. Rev. Math. Phys. **12**, 475–560 (2000)

[14] Buchholz, D., Lechner, G.: Modular nuclearity and localization. Ann. Henri Poincaré **5**, 1065–1080 (2004)

[15] Buchholz, D., Mund, J., Summers, S.J.: Covariant and quasi-covariant quantum dynamics in Robertson-Walker spacetimes. Class. Quantum Grav. **19**, 6417–6434 (2002)

[16] Budic, R., Isenberg, J., Lindblom, L., Yasskin, P.B.: On the determination of Cauchy surfaces from intrinsic properties. Commun. Math. Phys. **61**, 87–95 (1978)

[17] Connes, A., Rovelli, C.: von Neumann algebra automorphisms and time-thermodynamics relation in generally covariant quantum theories. Class. Quantum Grav. **11**, 2899–2917 (1994)

[18] Dappiaggi, C., Fredenhagen, K., Pinamonti, N.: Stable cosmological models driven by a free quantum scalar field. Phys. Rev. **D77**, 104015 (2008)

[19] Dappiaggi, C., Hack, T.P., Pinamonti, N.: The extended algebra of observables for Dirac fields and the trace anomaly of their stress-energy tensor. Rev. Math. Phys. **21**, 1241–1312 (2009)

[20] Degner, A., Verch, R.: Cosmological particle creation in states of low energy. J. Math. Phys. **51**, 022302 (2010)

[21] Dikranjan, D., Tholen, W.: Categorical structure of closure operators, *Mathematics and its Applications*, vol. 346. Kluwer Academic Publishers Group, Dordrecht (1995)

[22] Dimock, J.: Algebras of local observables on a manifold. Commun. Math. Phys. **77**, 219–228 (1980)

[23] Ferguson, M.: Dynamical locality of the nonminimally coupled scalar field and enlarged algebra of Wick polynomials. arXiv:1203.2151

[24] Fewster, C.J.: Endomorphisms and automorphisms of locally covariant quantum field theories. arXiv:1201.3295

[25] Fewster, C.J.: Quantum energy inequalities and local covariance. II. Categorical formulation. Gen. Relativity Gravitation **39**, 1855–1890 (2007)

[26] Fewster, C.J.: On the notion of 'the same physics in all spacetimes'. In: F. Finster, O. Müller, M. Nardmann, J. Tolksdorf, E. Zeidler (eds.) Quantum Field Theory and Gravity. Conceptual and mathematical advances in the search for a unified framework. Birkhäuser (2012). arXiv:1105.6202

[27] Fewster, C.J., Pfenning, M.J.: Quantum energy inequalities and local covariance. I: Globally hyperbolic spacetimes. J. Math. Phys. **47**, 082303 (2006)

[28] Fewster, C.J., Verch, R.: Dynamical locality of the free scalar field. arXiv:1109.6732, to appear in Annales H. Poincaré

[29] Fulling, S.A., Narcowich, F.J., Wald, R.M.: Singularity structure of the two-point function in quantum field theory in curved spacetime. II. Ann. Physics **136**, 243–272 (1981)

[30] Haag, R.: Local Quantum Physics: Fields, Particles, Algebras. Springer-Verlag, Berlin (1992)

[31] Hawking, S.W., Ellis, G.F.R.: The Large Scale Structure of Space-Time. Cambridge University Press, London (1973)

[32] Hollands, S., Wald, R.M.: Local Wick polynomials and time ordered products of quantum fields in curved spacetime. Commun. Math. Phys. **223**, 289–326 (2001)

[33] Hollands, S., Wald, R.M.: Existence of local covariant time ordered products of quantum fields in curved spacetime. Commun. Math. Phys. **231**, 309–345 (2002)

[34] Kay, B.S.: Linear spin-zero quantum fields in external gravitational and scalar fields. I. A one particle structure for the stationary case. Commun. Math. Phys. **62**, 55–70 (1978)

[35] Kay, B.S.: Casimir effect in quantum field theory. Phys. Rev. **D20**, 3052–3062 (1979)

[36] Kay, B.S.: The principle of locality and quantum field theory on (non-globally hyperbolic) curved spacetimes. Rev. Math. Phys. (Special Issue), 167–195 (1992)

[37] Kuckert, B.: Localization regions of local observables. Commun. Math. Phys. **215**, 197–216 (2000)

[38] Landau, L.J.: A note on extended locality. Commun. Math. Phys. **13**, 246–253 (1969)

[39] Landau, L.J.: On local functions of fields. Commun. Math. Phys. **39**, 49–62 (1974)

[40] Lee, J.M.: Introduction to smooth manifolds, *Graduate Texts in Mathematics*, vol. 218. Springer-Verlag, New York (2003)

[41] Mac Lane, S.: Categories for the Working Mathematician, 2nd edn. Springer-Verlag, New York (1998)

[42] Müllner, D.: Orientation reversal of manifolds. Algebr. Geom. Topol. **9**, 2361–2390 (2009)

[43] Nomizu, K., Ozeki, H.: The existence of complete Riemannian metrics. Proc. Amer. Math. Soc. **12**, 889–891 (1961)

[44] O'Neill, B.: Semi-Riemannian Geometry. Academic Press, New York (1983)

[45] Penrose, R.: Techniques of differential topology in relativity. Society for Industrial and Applied Mathematics, Philadelphia, Pa. (1972). Conference Board of the Mathematical Sciences Regional Conference Series in Applied Mathematics, No. 7

[46] Ruzzi, G.: Homotopy of posets, net-cohomology and superselection sectors in globally hyperbolic space-times. Rev. Math. Phys. **17**, 1021–1070 (2005)

[47] Sanders, K.: On the Reeh-Schlieder property in curved spacetime. Commun. Math. Phys. **288**, 271–285 (2009)

[48] Sanders, K.: The locally covariant Dirac field. Rev. Math. Phys. **22**, 381–430 (2010)

[49] Schoch, A.: On the simplicity of Haag fields. Int. J. Theor. Phys. **1**, 107–113 (1968)

[50] Verch, R.: A spin-statistics theorem for quantum fields on curved spacetime manifolds in a generally covariant framework. Commun. Math. Phys. **223**, 261–288 (2001)

[51] Verch, R.: Local covariance, renormalization ambiguity, and local thermal equilibrium in cosmology. In: F. Finster, O. Müller, M. Nardmann, J. Tolksdorf, E. Zeidler (eds.) Quantum Field Theory and Gravity. Conceptual and mathematical advances in the search for a unified framework. Birkhäuser (2012). arXiv:1105.6249

[52] Wald, R.M.: General Relativity. University of Chicago Press, Chicago (1984)

[53] Wiesbrock, H.W.: Modular intersections of von Neumann algebras in quantum field theory. Commun. Math. Phys. **193**, 269–285 (1998)